OM protein - protein search, using sw model

Run on: April 15, 2004, 22:18:54 ; Search time 366 Seconds
(without alignments)
1895.818 Million cell updates/sec

Title: US-09-522-753-5
Perfect score: 13215
Sequence: 1 MSGSTQLVAQTWRATEPRYP.....WDEEPKPLLCSQYETLSDSE 2517

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1124875 seqs, 275673149 residues

Total number of hits satisfying chosen parameters: 1124875

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	13145	99.5	2517	12	US-10-087-192-654	Sequence 654, App
2	12978	98.2	2507	10	US-09-819-104A-2	Sequence 2, Appli
3	10796.5	81.7	2462	10	US-09-819-104A-5	Sequence 5, Appli
4	7871	59.6	1495	14	US-10-351-750-1	Sequence 1, Appli
5	6685	50.6	1585	12	US-10-087-192-651	Sequence 651, App
6	4611	34.9	876	14	US-10-146-473-54	Sequence 54, Appl
7	4165.5	31.5	2440	15	US-10-341-434-236	Sequence 236, App
8	569.5	4.3	3664	12	US-10-263-929-143	Sequence 143, App
9	569.5	4.3	3664	14	US-10-177-293-423	Sequence 423, App
10	547	4.1	3551	12	US-10-263-929-144	Sequence 144, App
11	518	3.9	2665	9	US-09-864-761-34248	Sequence 34248, A
12	485.5	3.7	2803	12	US-10-415-187-5	Sequence 5, Appli
13	472.5	3.6	5262	15	US-10-051-874-167	Sequence 167, App
14	470.5	3.6	4952	15	US-10-051-874-56	Sequence 56, Appl
15	470.5	3.6	5159	15	US-10-085-198-112	Sequence 112, App
16	470.5	3.6	5262	15	US-10-051-874-165	Sequence 165, App
17	469.5	3.6	2701	14	US-10-171-311-83	Sequence 83, Appl
18	466.5	3.5	5560	12	US-10-263-929-142	Sequence 142, App
19	452	3.4	2971	14	US-10-146-473-50	Sequence 50, Appl
20	451	3.4	5008	15	US-10-051-874-166	Sequence 166, App
21	436	3.3	1151	12	US-09-825-751A-79	Sequence 79, Appl
22	436	3.3	4019	9	US-09-738-973-425	Sequence 425, App
23	436	3.3	4019	9	US-09-854-133-425	Sequence 425, App
24	436	3.3	4019	14	US-10-144-649A-425	Sequence 425, App
25	429.5	3.3	2429	12	US-10-362-892-8	Sequence 8, Appli
26	429.5	3.3	2429	15	US-10-288-798-8	Sequence 8, Appli
27	428	3.2	1363	13	US-10-124-557-52	Sequence 52, Appl
28	428	3.2	1404	12	US-09-802-207-30	Sequence 30, Appl
29	428	3.2	1404	13	US-10-124-557-2	Sequence 2, Appli
30	428	3.2	1404	13	US-10-124-557-62	Sequence 62, Appl
31	428	3.2	2063	9	US-09-735-367B-2	Sequence 2, Appli
32	425.5	3.2	1320	13	US-10-124-557-46	Sequence 46, Appl
33	425.5	3.2	1320	13	US-10-124-557-60	Sequence 60, Appl
34	425.5	3.2	1361	13	US-10-124-557-40	Sequence 40, Appl
35	425.5	3.2	2545	12	US-10-042-865-12	Sequence 12, Appl
36	424.5	3.2	2429	12	US-10-377-035-17	Sequence 17, Appl
37	424.5	3.2	2432	12	US-10-112-944-359	Sequence 359, App
38	423	3.2	1140	13	US-10-124-557-104	Sequence 104, App
39	421.5	3.2	1313	13	US-10-124-557-142	Sequence 142, App
40	421.5	3.2	1354	13	US-10-124-557-48	Sequence 48, Appl
41	419.5	3.2	1314	13	US-10-124-557-50	Sequence 50, Appl
42	418	3.2	1828	12	US-10-221-625-15	Sequence 15, Appl
43	417	3.2	2005	9	US-09-735-367B-3	Sequence 3, Appli
44	416.5	3.2	1049	13	US-10-124-557-58	Sequence 58, Appl
45	414	3.1	2092	12	US-10-042-865-79	Sequence 79, Appl
46	414	3.1	2092	12	US-10-377-035-18	Sequence 18, Appl
47	414	3.1	2137	12	US-10-042-865-81	Sequence 81, Appl
48	410	3.1	1270	13	US-10-124-557-44	Sequence 44, Appl
49	409	3.1	19652	15	US-10-084-846A-7	Sequence 7, Appli
50	408.5	3.1	1038	13	US-10-124-557-74	Sequence 74, Appl
51	408.5	3.1	1311	13	US-10-124-557-42	Sequence 42, Appl
52	406.5	3.1	2364	12	US-10-205-331-66	Sequence 66, Appl
53	406	3.1	2382	14	US-10-196-935A-2	Sequence 2, Appli
54	405	3.1	2382	12	US-10-336-472-230	Sequence 230, App
55	405	3.1	2382	15	US-10-052-648A-40	Sequence 40, Appl
56	400.5	3.0	941	13	US-10-124-557-14	Sequence 14, Appl
57	400.5	3.0	1022	13	US-10-124-557-84	Sequence 84, Appl
58	400.5	3.0	1464	14	US-10-216-705-21	Sequence 21, Appl
59	398	3.0	1000	14	US-10-128-714-3305	Sequence 3305, Ap
60	396.5	3.0	1464	15	US-10-291-265-243	Sequence 243, App
61	394	3.0	2783	9	US-09-816-669A-14	Sequence 14, Appl
62	394	3.0	19723	15	US-10-084-846A-5	Sequence 5, Appli
63	392.5	3.0	1464	12	US-09-918-715-261	Sequence 261, App
64	392.5	3.0	1464	14	US-10-060-036-159	Sequence 159, App
65	392.5	3.0	1464	14	US-10-171-311-36	Sequence 36, Appl
66	392.5	3.0	1464	14	US-10-149-352-2	Sequence 2, Appli
67	392.5	3.0	1464	14	US-10-177-293-65	Sequence 65, Appl
68	392.5	3.0	1464	14	US-10-301-822-28	Sequence 28, Appl
69	391.5	3.0	1633	14	US-10-029-386-33090	Sequence 33090, A
70	388.5	2.9	2058	15	US-10-052-648A-20	Sequence 20, Appl
71	386	2.9	3252	15	US-10-210-130-36	Sequence 36, Appl
72	383.5	2.9	19662	15	US-10-084-846A-6	Sequence 6, Appli
73	381.5	2.9	19725	15	US-10-084-846A-4	Sequence 4, Appli
74	380.5	2.9	4723	15	US-10-359-012-8	Sequence 8, Appli
75	380	2.9	2231	15	US-10-379-381-5	Sequence 5, Appli
76	379.5	2.9	1435	12	US-10-276-774-2178	Sequence 2178, Ap
77	377.5	2.9	4823	15	US-10-051-874-169	Sequence 169, App
78	377	2.9	2380	12	US-10-333-314-18	Sequence 18, Appl
79	377	2.9	3208	15	US-10-210-130-38	Sequence 38, Appl
80	376.5	2.8	3268	15	US-10-379-381-2	Sequence 2, Appli
81	375.5	2.8	1390	12	US-10-092-900A-224	Sequence 224, App
82	374.5	2.8	3186	15	US-10-210-130-34	Sequence 34, Appl
83	374	2.8	19695	15	US-10-084-846A-3	Sequence 3, Appli
84	371	2.8	1463	15	US-10-402-089-2	Sequence 2, Appli
85	371	2.8	1463	15	US-10-402-072A-2	Sequence 2, Appli
86	369.5	2.8	3262	15	US-10-379-381-4	Sequence 4, Appli
87	367	2.8	19608	15	US-10-084-846A-8	Sequence 8, Appli
88	366	2.8	1449	15	US-10-402-089-8	Sequence 8, Appli

89	366	2.8	1449	15	US-10-402-072A-8	Sequence 8, Appli
90	365	2.8	960	15	US-10-342-331-5	Sequence 5, Appli
91	364	2.8	1426	12	US-10-664-859-15	Sequence 15, Appl
92	364	2.8	1426	12	US-09-915-543-15	Sequence 15, Appl
93	364	2.8	1426	14	US-10-322-579-15	Sequence 15, Appl
94	358	2.7	1618	9	US-09-963-875-1	Sequence 1, Appli
95	358	2.7	1618	12	US-09-731-255-2	Sequence 2, Appli
96	358	2.7	1618	14	US-10-136-891-2	Sequence 2, Appli
97	358	2.7	1618	14	US-10-120-687-1	Sequence 1, Appli
98	355	2.7	2093	14	US-10-032-585-7665	Sequence 7665, Ap
99	354	2.7	3507	15	US-10-369-493-5784	Sequence 5784, Ap
100	351	2.7	1341	14	US-10-058-124-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1
US-10-087-192-654
; Sequence 654, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 654
; LENGTH: 2517
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-087-192-654

Query Match 99.5%; Score 13145; DB 12; Length 2517;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 2505; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

```
Qy          1 MSGSTQLVAQTWRATEPRYPPHSLSYPVQIARTHTDVGLLEYQHHSRDYASHLSPGSIIQ 60
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSGSTQPVAQTWRATEPRYPPHSLSYPVQIARTHTDVGLLEYQHHSRDYASHLSPGSIIQ 60

Qy         61 PQRRRPSLLSEFQPGNERSQELHLRPESHSYLPELGKSEMEFIESKRPRLELLPDPLLRP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PQRRRPSLLSEFQPGNERSQELHLRPESHSYLPELGKSEMEFIESKRPRLELLPDPLLRP 120

Qy        121 SPLLATGQPAGSEDLTKDRSLTGKLEPVSPPSPPHTDPELELVPPRLSKEELIQNMDRVD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SPLLATGQPAGSEDLTKDRSLTGKLEPVSPPSPPHTDPELELVPPRLSKEELIQNMDRVD 180

Qy        181 REITMVEQQISKLKKKQQQLEEEAAKPPEPEKPVSPPPIESKHRSLVQIIYDENRKKAEA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 REITMVEQQISKLKKKQQQLEEEAAKPPEPEKPVSPPPIESKHRSLVQIIYDENRKKAEA 240

Qy        241 AHRILEGLGPQVELPLYNQPSDTRQYHENIKINQAMRKKLILYFKRRNHARKQWKQKFCQ 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db        241 AHRILEGLGPQVELPLYNQPSDTRQYHENIKINQAMRKKLILYFKRRNHARKQWEQKFCQ 300

Qy        301 RYDQLMEALEKKVERIENNPRRRAKESKVREYYEKQFPEIRKQRELQERMQSRVGQRGSG 360
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 RYDQLMEAWEKKVERIENNPRRRAKESKVREYYEKQFPEIRKQRELQERMQSRVGQRGSG 360

Qy        361 LSMSAARSEHEVSEIIDGLSEQENLEKQMRQLAVIPPMLYDADQQRIKFINMNGLMADPM 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LSMSAARSEHEVSEIIDGLSEQENLEKQMRQLAVIPPMLYDADQQRIKFINMNGLMADPM 420

Qy        421 KVYKDRQVMNMWSEQEKETFREKFMQHPKNFGLIASFLERKTVAECVLYYYLTKKNENYK 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 KVYKDRQVMNMWSEQEKETFREKFMQHPKNFGLIASFLERKTVAECVLYYYLTKKNENYK 480

Qy        481 SLVRRSYRRRGKSQQQQQQQQQQQQQQQQQQPMPRSSQEEKDEKEKEKEAEKEEEKPEVEN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 SLVRRSYRRRGKSQQQQQQQQQQQQQQQQQQPMPRSSQEEKDEKEKEKEAEKEEEKPEVEN 540

Qy        541 DKEDLLKEKTDDTSGEDNDEKEAVASKGRKTANSQGRRKGRITRSMANEANSEEAITPQQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 DKEDLLKEKTDDTSGEDNDEKEAVASKGRKTANSQGRRKGRITRSMANEANSEEAITPQQ 600

Qy        601 SAELASMELNESSRWTEEEMETAKKGLLEHGRNWSAIARMVGSKTVSQCKNFYFNYKKRQ 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 SAELASMELNESSRWTEEEMETAKKGLLEHGRNWSAIARMVGSKTVSQCKNFYFNYKKRQ 660

Qy        661 NLDEILQQHKLKMEKERNARRKKKKAPAAASEEAAFPPVVEDEEMEASGVSGNEEEMVEE 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 NLDEILQQHKLKMEKERNARRKKKKAPAAASEEAAFPPVVEDEEMEASGVSGNEEEMVEE 720

Qy        721 AEALHASGNEVPRGECSGPATVNNSSDTESIPSPHTEAAKDTGQNGPKPPATLGADGPPP 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AEALHASGNEVPRGECSGPATVNNSSDTESIPSPHTEAAKDTGQNGPKPPATLGADGPPP 780

Qy        781 GPPTPPRRTSRAPIEPTPASEATGAPTPPPAPPSPSAPPPVVPKEEKEEETAAAPPVEEG 840
              ||||||     || ||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GPPTPPPEDIPAPTEPTPASEATGAPTPPPAPPSPSAPPPVVPKEEKEEETAAAPPVEEG 840

Qy        841 EEQKPPAAEELAVDTGKAEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGS 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 EEQKPPAAEELAVDTGKAEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGS 900

Qy        901 GRATTAKSSGAPQDSDSSATCSADEVDEAEGGDKNRLLSPRPSLLTPTGDPRANASPQKP 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GRATTAKSSGAPQDSDSSATCSADEVDEAEGGDKNRLLSPRPSLLTPTGDPRANASPQKP 960

Qy        961 LDLKQLKQRAAAIPPIQVTKVHEPPREDAAPTKPAPPAPPPPQNLQPESDAPQQPGSSPR 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 LDLKQLKQRAAAIPPIQVTKVHEPPREDAAPTKPAPPAPPPPQNLQPESDAPQQPGSSPR 1020

Qy       1021 GKSRSPAPPADKEAFAAEAQKLPGDPPCWTSGLPFPVPPREVIKASPHAPDPSAFSYAPP 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GKSRSPAPPADKEAFAAEAQKLPGDPPCWTSGLPFPVPPREVIKASPHAPDPSAFSYAPP 1080

Qy       1081 GHPLPLGLHDTARPVLPRPPTISNPPPLISSAKHPSVLERQIGAISQGMSVQLHVPYSEH 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 GHPLPLGLHDTARPVLPRPPTISNPPPLISSAKHPSVLERQIGAISQGMSVQLHVPYSEH 1140

Qy       1141 AKAPVGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPESLGVPTAQEASVLRGTAL 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 AKAPVGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPESLGVPTAQEASVLRGTAL 1200

Qy       1201 GSVPGGSITKGIPSTRVPSDSAITYRGSITHGTPADVLYKGTITRIIGEDSPSRLDRGRE 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GSVPGGSITKGIPSTRVPSDSAITYRGSITHGTPADVLYKGTITRIIGEDSPSRLDRGRE 1260

Qy       1261 DSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPHETAAPKRTYDMMEGRVGR 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 DSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPHETAAPKRTYDMMEGRVGR 1320

Qy       1321 AISSASIEGLMGRAIPPERHSPHHLKEQHHIRGSITQGIPRSYVEAQEDYLRREAKLLKR 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 AISSASIEGLMGRAIPPERHSPHHLKEQHHIRGSITQGIPRSYVEAQEDYLRREAKLLKR 1380

Qy       1381 EGTPPPPPPSRDLTEAYKTQALGPLKLKPAHEGLVATVKEAGRSIHEIPREELRHTPELP 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 EGTPPPPPPSRDLTEAYKTQALGPLKLKPAHEGLVATVKEAGRSIHEIPREELRHTPELP 1440

Qy       1441 LAPRPLKEGSITQGTPLKYDTGASTTGSKKHDVRSLIGSPGRTFPPVHPLDVMADARALE 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 LAPRPLKEGSITQGTPLKYDTGASTTGSKKHDVRSLIGSPGRTFPPVHPLDVMADARALE 1500

Qy       1501 RACYEESLKSRPGTASSSGGSIARGAPVIVPELGKPRQSPLTYEDHGAPFAGHLPRGSPV 1560
```

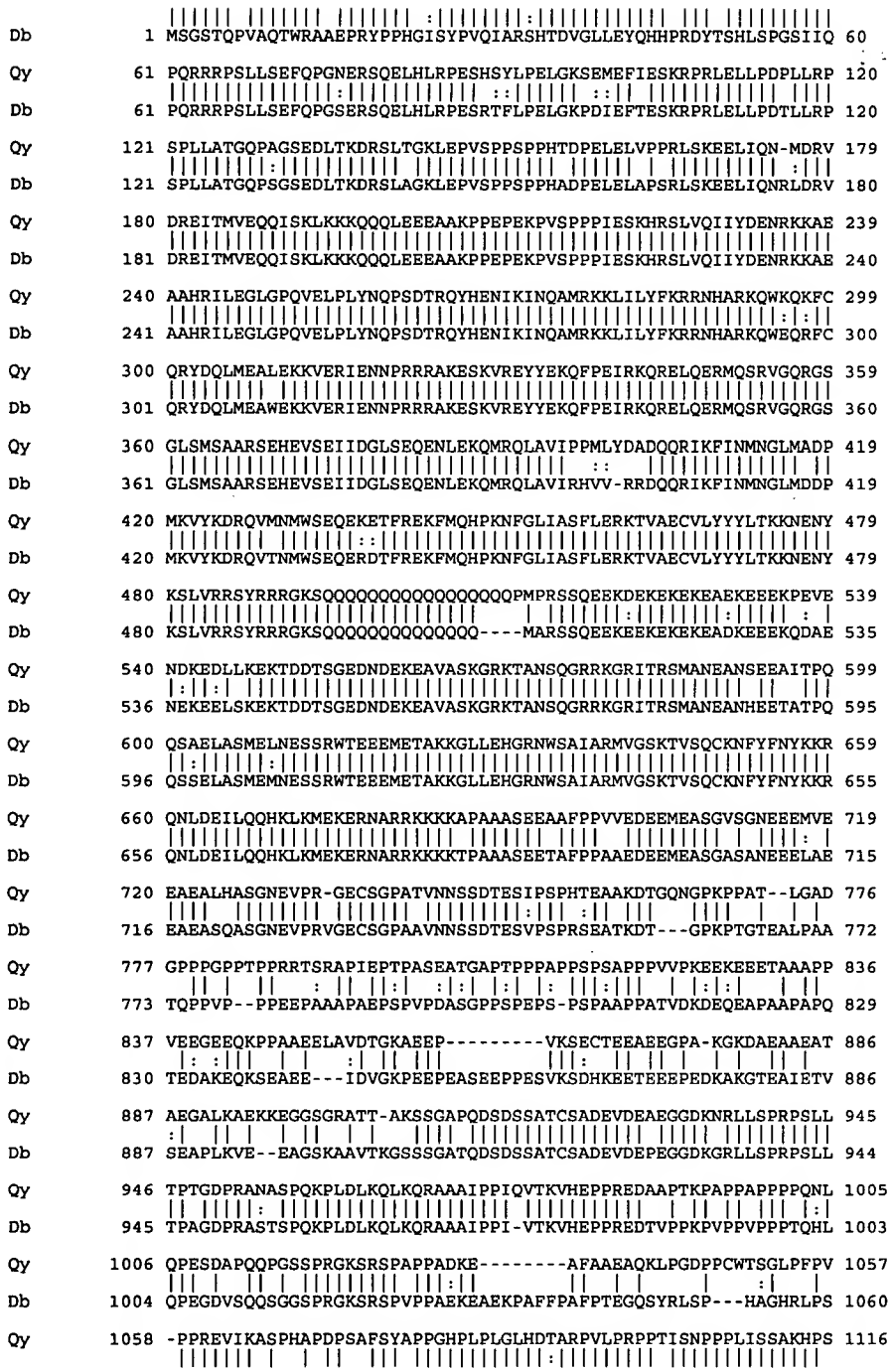
```
Db       1501 RACYEESLKSRPGTASSSGGSIARGAPVIVPELGKPRQSPLTYEDHGAPFAGHLPRGSPV 1560

Qy       1561 TMREPTPRLQEGSLSSSKASQDRKLTSTPREIAKSPHSTVPEHHPHPISPYEHLLRGVSG 1620
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 TTREPTPRLQEGSLSSSKASQDRKLTSTPREIAKSPHSTVPEHHPHPISPYEHLLRGVSG 1620

Qy       1621 VDLYRSHIPLAFDPTSIPRGIPLDAAAAYYLPRHLAPNPTYPHLYPPYLIRGYPDTAALE 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 VDLYRSHIPLAFDPTSIPRGIPLDAAAAYYLPRHLAPNPTYPHLYPPYLIRGYPDTAALE 1680

Qy       1681 NRQTIINDYITSQQMHHNTATAMAQRADMLRGLSPRESSLALNYAAGPRGIIDLSQVPHL 1740
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db       1681 NRQTIINDYITSQQMHHNAATAMAQRADMLRGLSPRESSLALNYAAGPRGIIDLSQVPHL 1740

Qy       1741 PVLVPPTPGTPATAMDRLAYLPTAPQPFSSRHSSSPLSPGGPTHLTKPTTTSSSERERDR 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 PVLVPPTPGTPATAMDRLAYLPTAPQPFSSRHSSSPLSPGGPTHLTKPTTTSSSERERDR 1800

Qy       1801 DRERDRDREREKSILTSTTTVEHAPIWRPGTEQSSGSSGSSGGGGGSSSRPASHSHAHQH 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 DRERDRDREREKSILTSTTTVEHAPIWRPGTEQSSGSSGSSGGGGGSSSRPASHSHAHQH 1860

Qy       1861 SPISPRTQDALQQRPSVLHNTGMKGIITAVEPSKPTVLRSTSTSSPVRPAATFPPATHCP 1920
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db       1861 SPISPRTQDALQQRPSVLHNTGMKGIITAVEPSTPTVLRSTSTSSPVRPAATFPPATHCP 1920

Qy       1921 LGGTLDGVYPTLMEPVLLPKEAPRVARPERPRADTGHAFLAKPPARSGLEPASSPSKGSE 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 LGGTLDGVYPTLMEPVLLPKEAPRVARPERPRADTGHAFLAKPPARSGLEPASSPSKGSE 1980

Qy       1981 PRPLVPPVSGHATIARTPAKNLAPHHASPDPPAPPASASDPHREKTQSKPFSIQELELRS 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 PRPLVPPVSGHATIARTPAKNLAPHHASPDPPAPPASASDPHREKTQSKPFSIQELELRS 2040

Qy       2041 LGYHGSSYSPEGVEPVSPVSSPSLTHDKGLPKHLEELDKSHLEGELRPKQPGPVKLGGEA 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 LGYHGSSYSPEGVEPVSPVSSPSLTHDKGLPKHLEELDKSHLEGELRPKQPGPVKLGGEA 2100

Qy       2101 AHLPHLRPLPESQPSSSPLLQTAPGVKGHQRVVTLAQHISEVITQDYTRHHPQQLSAPLP 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 AHLPHLRPLPESQPSSSPLLQTAPGVKGHQRVVTLAQHISEVITQDYTRHHPQQLSAPLP 2160

Qy       2161 APLYSFPGASCPVLDLRRPPSDLYLPPPDHGAPARGSPHSEGGKRSPEPNKTSVLGGGED 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 APLYSFPGASCPVLDLRRPPSDLYLPPPDHGAPARGSPHSEGGKRSPEPNKTSVLGGGED 2220

Qy       2221 GIEPVSPPEGMTEPGHSRSAVYPLLYRDGEQTEPSRMGSKSPGNTSQPPAFFSKLTESNS 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 GIEPVSPPEGMTEPGHSRSAVYPLLYRDGEQTEPSRMGSKSPGNTSQPPAFFSKLTESNS 2280

Qy       2281 AMVKSKKQEINKKLNTHNRNEPEYNISQPGTEIFNMPAITGTGLMTYRSQAVQEHASTNM 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 AMVKSKKQEINKKLNTHNRNEPEYNISQPGTEIFNMPAITGTGLMTYRSQAVQEHASTNM 2340

Qy       2341 GLEAIIRKALMGKYDQWEESPPLSANAFNPLNASASLPAAMPITAADGRSDHTLTSPGGG 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 GLEAIIRKALMGKYDQWEESPPLSANAFNPLNASASLPAAMPITAADGRSDHTLTSPGGG 2400

Qy       2401 GKAKVSGRPSSRKAKSPAPGLASGDRPPSVSSVHSEGDCNRRTPLTNRVWEDRPSSAGST 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 GKAKVSGRPSSRKAKSPAPGLASGDRPPSVSSVHSEGDCNRRTPLTNRVWEDRPSSAGST 2460

Qy       2461 PFPYNPLIMRLQAGVMASPPPPGLPAGSGPLAGPHHAWDEEPKPLLCSQYETLSDSE 2517
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 PFPYNPLIMRLQAGVMASPPPPGLPAGSGPLAGPHHAWDEEPKPLLCSQYETLSDSE 2517


RESULT 2
US-09-819-104A-2
; Sequence 2, Application US/09819104A
; Publication No. US20030027137A1
; GENERAL INFORMATION:
;  APPLICANT: Chen, J. Don
;  TITLE OF INVENTION: NOVEL NUCLEAR RECEPTOR COREPRESSOR MOLECULES
;  TITLE OF INVENTION:  AND USES THEREFOR
;  FILE REFERENCE: UMG-030
;  CURRENT APPLICATION NUMBER: US/09/819,104A
;  CURRENT FILING DATE:  2001-03-27
;  PRIOR APPLICATION NUMBER: 60/193,138
;  PRIOR FILING DATE: 2000-03-29
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 2507
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-819-104A-2

  Query Match             98.2%;  Score 12978;  DB 10;  Length 2507;
  Best Local Similarity   98.3%;  Pred. No. 0;
  Matches 2483;  Conservative    3;  Mismatches   13;  Indels   26;  Gaps    3;

Qy          1 MSGSTQLVAQTWRATEPRYPPHSLSYPVQIARTHTDVGLLEYQHHSRDYASHLSPGSIIQ 60
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSGSTQPVAQTWRATEPRYPPHSLSYPVQIARTHTDVGLLEYQHHSRDYASHLSPGSIIQ 60

Qy         61 PQRRRPSLLSEFQPGNERSQELHLRPESHSYLPELGKSEMEFIESKRPRLELLPDPLLRP 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PQRRRPSLLSEFQPGNERSQELHLRPESHSYLPELGKSEMEFIESKRPRLELLPDPLLRP 120

Qy        121 SPLLATGQPAGSEDLTKDRSLTGKLEPVSPPSPPHTDPELELVPPRLSKEELIQNMDRVD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SPLLATGQPAGSEDLTKDRSLTGKLEPVSPPSPPHTDPELELVPPRLSKEELIQNMDRVD 180

Qy        181 REITMVEQQISKLKKKQQQLEEEAAKPPEPEKPVSPPPIESKHRSLVQIIYDENRKKAEA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 REITMVEQQISKLKKKQQQLEEEAAKPPEPEKPVSPPPIESKHRSLVQIIYDENRKKAEA 240

Qy        241 AHRILEGLGPQVELPLYNQPSDTRQYHENIKINQAMRKKLILYFKRRNHARKQWKQKFCQ 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db        241 AHRILEGLGPQVELPLYNQPSDTRQYHENIKINQAMRKKLILYFKRRNHARKQWEQKFCQ 300

Qy        301 RYDQLMEALEKKVERIENNPRRRAKESKVREYYEKQFPEIRKQRELQERMQSRVGQRGSG 360
              |||||||| ||||||||||||||||||||||||||||||||||||||| ||||||||
Db        301 RYDQLMEAWEKKVERIENNPRRRAKESKVREYYEKQFPEIRKQRELQERMQ-RVGQRGSG 359

Qy        361 LSMSAARSEHEVSEIIDGLSEQENLEKQMRQLAVIPPMLYDADQQRIKFINMNGLMADPM 420
              |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 LSMSPARSEHEVSEIIDGLSEQENLEKQMRQLAVIPPMLYDADQQRIKFINMNGLMADPM 419

Qy        421 KVYKDRQVMNMWSEQEKETFREKFMQHPKNFGLIASFLERKTVAECVLYYYLTKKNENYK 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 KVYKDRQVMNMWSEQEKETFREKFMQHPKNFGLIASFLERKTVAECVLYYYLTKKNENYK 479

Qy        481 SLVRRSYRRRGKSQQQQQQQQQQQQQQQQQPMPRSSQEEKDEKEKEKEAEKEEEKPEVEN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 SLVRRSYRRRGKSQQQQQQQQQQQQQQQQQPMPRSSQEEKDEKEKEKEAEKEEEKPEVEN 539

Qy        541 DKEDLLKEKTDDTSGEDNDEKEAVASKGRKTANSQGRRKGRITRSMANEANSEEAITPQQ 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        540 DKEDLLKEKTDDTSGEDNDEKEAVASKGRKTANSQGRRKGRITRSMANEANSEEAITPQQ 599

Qy        601 SAELASMELNESSRWTEEEMETAKKGLLEHGRNWSAIARMVGSKTVSQCKNFYFNYKKRQ 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        600 SAELASMELNESSRWTEEEMETAKKGLLEHGRNWSAIARMVGSKTVSQCKNFYFNYKKRQ 659

Qy        661 NLDEILQQHKLKMEKERNARRKKKKAPAAASEEAAFPPVVEDEEMEASGVSGNEEEMVEE 720
              |||||||||||||||||||||||||||||||||||||||||||||||||:|||||||||
Db        660 NLDEILQQHKLKMEKERNARRKKKKAPAAASEEAAFPPVVEDEEMEASGVTGNEEEMVEE 719

Qy        721 AEALHASGNEVPRGECSGPATVNNSSDTESIPSPHTEAAKDTGQNGPKPPATLGADGPPP 780
              ||                 |||||||||||||||||||||||||||||||||||||||
Db        720 AE-----------------ATVNNSSDTESIPSPHTEAAKDTGQNGPKPPATLGADGPPP 762

Qy        781 GPPTPPRRTSRAPIEPTPASEATGAPTPPPAPPSPSAPPPVVPKEEKEEETAAAPPVEEG 840
```

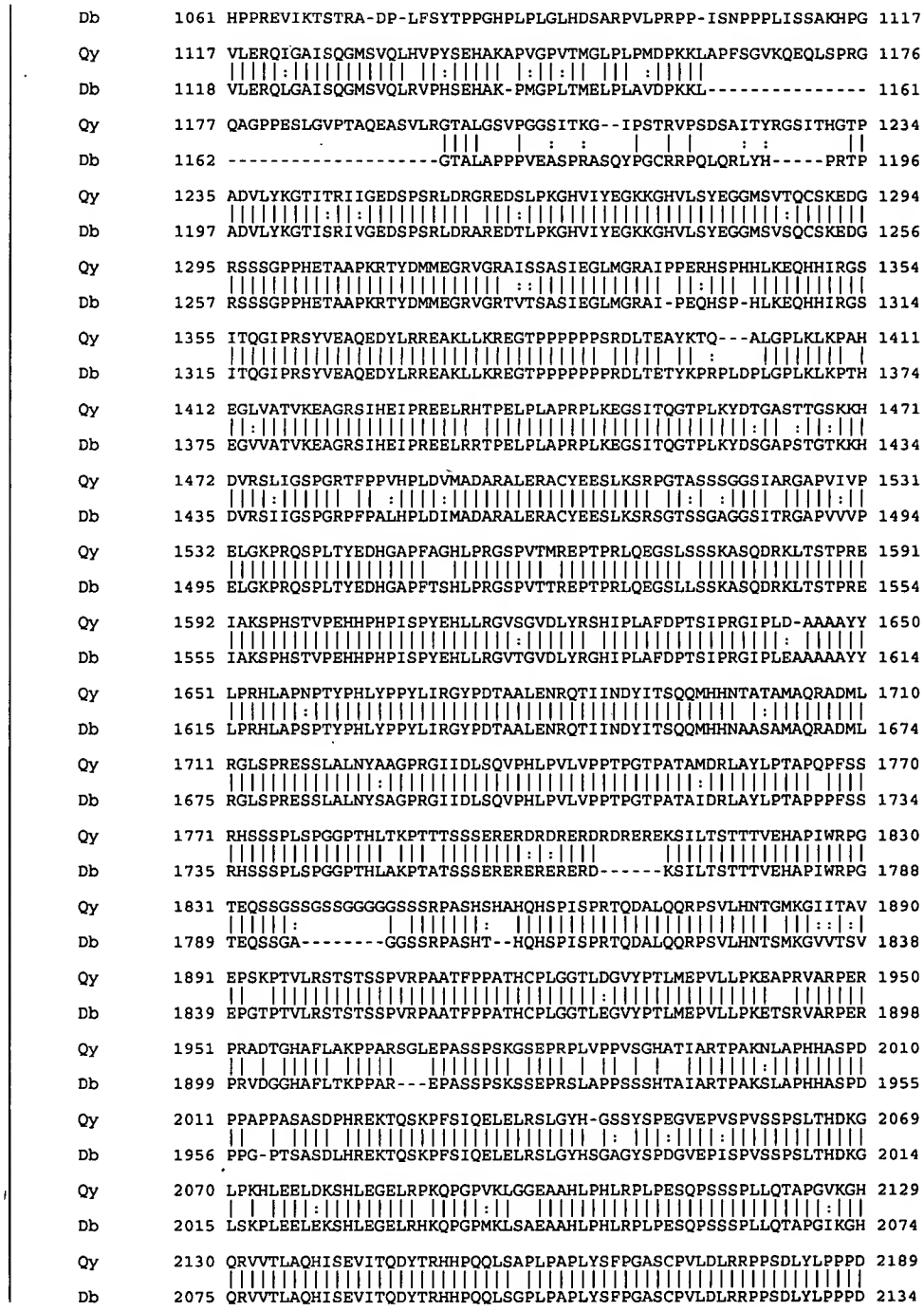
```
Db        763 GPPTPPPEDIPAPTESTPASEATLAPTPPPAPPFPSSPPPVVPKEEKEEETAAAPPVEEG 822

Qy        841 EEQKPPAAEELAVDTGKAEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGS 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        823 EEQKPPAAEELAVDTGKAEEPVKSECTEEAEEGPAKGKDAEAAEATAERALKAEKKEGGS 882

Qy        901 GRATTAKSSGAPQDSDSSATCSADEVDEAEGGDKNRLLSPRPSLLTPTGDPRANASPQKP 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        883 GRATTAKSSGAPQDSDSSATCSADEVDEAEGGDKNRLLSPRPSLLTPTGDPRANASPQKP 942

Qy        961 LDLKQLKQRAAAIPPIQVTKVHEPPREDAAPTKPAPPAPPPPQNLQPESDAPQQPGSSPR 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        943 LDLKQLKQRAAAIPPIQVTKVHEPPREDAAPTKPAPPAPPPPQNLQPESDAPQQPGSSPR 1002

Qy       1021 GKSRSPAPPADKE--------AFAAEAQKLPGDPPCWTSGLPFPVPPREVIKASPHAPDP 1072
              |||||||||||||        ||||||||||||||||||||||||||||||||||||||||
Db       1003 GKSRSPAPPADKEAEKPVFFPAFAAEAQKLPGDPPCWTSGLPFPVPPREVIKASPHAPDP 1062

Qy       1073 SAFSYAPPGHPLPLGLHDTARPVLPRPPTISNPPPLISSAKHPSVLERQIGAISQGMSVQ 1132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1063 SAFSYAPPGHPLPLGLHDTARPVLPRPPTISNPPPLISSAKHPSVLERQIGAISQGMSVQ 1122

Qy       1133 LHVPYSEHAKAPVGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPESLGVPTAQEA 1192
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1123 LHVPYSEHAKAPVGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPESLGVPTAQEA 1182

Qy       1193 SVLRGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHGTPADVLYKGTITRIIGEDSP 1252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1183 SVLRGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHGTPADVLYKGTITRIIGEDSP 1242

Qy       1253 SRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPHETAAPKRTYD 1312
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1243 SRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPHETAAPKRTYD 1302

Qy       1313 MMEGRVGRAISSASIEGLMGRAIPPERHSPHHLKEQHHIRGSITQGIPRSYVEAQEDYLR 1372
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1303 MMEGRVGRAISSASIEGLMGRAIPPERHSPHHLKEQHHIRGSITQGIPRSYVEAQEDYLR 1362

Qy       1373 REAKLLKREGTPPPPPPSRDLTEAYKTQALGPLKLKPAHEGLVATVKEAGRSIHEIPREE 1432
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1363 REAKLLKREGTPPPPPPSRDLTEAYKTQALGPLKLKPAHEGLVATVKEAGRSIHEIPREE 1422

Qy       1433 LRHTPELPLAPRPLKEGSITQGTPLKYDTGASTTGSKKHDVRSLIGSPGRTFPPVHPLDV 1492
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1423 LRHTPELPLAPRPLKEGSITQGTPLKYDTGASTTGSKKHDVRSLIGSPGRTFPPVHPLDV 1482

Qy       1493 MADARALERACYEESLKSRPGTASSSGGSIARGAPVIVPELGKPRQSPLTYEDHGAPFAG 1552
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1483 MADARALERACYEESLKSRPGTASSSGGSIARGAPVIVPELGKPRQSPLTYEDHGAPFAG 1542

Qy       1553 HLPRGSPVTMREPTPRLQEGSLSSSKASQDRKLTSTPREIAKSPHSTVPEHHPHPISPYE 1612
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1543 HLPRGSPVTMREPTPRLQEGSLSSSKASQDRKLTSTPREIAKSPHSTVPEHHPHPISPYE 1602

Qy       1613 HLLRGVSGVDLYRSHIPLAFDPTSIPRGIPLDAAAAYYLPRHLAPNPTYPHLYPPYLIRG 1672
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1603 HLLRGVSGVDLYRSHIPLAFDPTSIPRGIPLDAAAAYYLPRHLAPNPTYPHLYPPYLIRG 1662

Qy       1673 YPDTAALENRQTIINDYITSQQMHHNTATAMAQRADMLRGLSPRESSLALNYAAGPRGII 1732
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1663 YPDTAALENRQTIINDYITSQQMHHNTATAMAQRADMLRGLSPRESSLALNYAAGPRGII 1722

Qy       1733 DLSQVPHLPVLVPPTPGTPATAMDRLAYLPTAPQPFSSRHSSSPLSPGGPTHLTKPTTTS 1792
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1723 DLSQVPHLPVLVPPTPGTPATAMDRLAYLPTAPQPFSSRHSSSPLSPGGPTHLTKPTTTS 1782

Qy       1793 SSERERDRDRERDRDREREKSILTSTTTVEHAPIWRPGTEQSSGSSGSSGGGGGSSSRPA 1852
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1783 SSERERDRDRERDRDREREKSILTSTTTVEHAPIWRPGTEQSSGSSGSSGGGGGSSSRPA 1842

Qy       1853 SHSHAHQHSPISPRTQDALQQRPSVLHNTGMKGIITAVEPSKPTVLRSTSTSSPVRPAAT 1912
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        1843 SHSHAHQHSPISPRTQDALQQRPSVLHNTGMKGIITAVEPSKPTVLRSTSTSSPVRPAAT 1902

Qy        1913 FPPATHCPLGGTLDGVYPTLMEPVLLPKEAPRVARPERPRADTGHAFLAKPPARSGLEPA 1972
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1903 FPPATHCPLGGTLDGVYPTLMEPVLLPKEAPRVARPERPRADTGHAFLAKPPARSGLEPA 1962

Qy        1973 SSPSKGSEPRPLVPPVSGHATIARTPAKNLAPHHASPDPPAPPASASDPHREKTQSKPFS 2032
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1963 SSPSKGSEPRPLVPPVSGHATIARTPAKNLAPHHASPDPPAPPASASDPHREKTQSKPFS 2022

Qy        2033 IQELELRSLGYHGSSYSPEGVEPVSPVSSPSLTHDKGLPKHLEELDKSHLEGELRPKQPG 2092
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2023 IQELELRSLGYHGSSYSPEGVEPVSPVSSPSLTHDKGLPKHLEELDKSHLEGELRPKQPG 2082

Qy        2093 PVKLGGEAAHLPHLRPLPESQPSSSPLLQTAPGVKGHQRVVTLAQHISEVITQDYTRHHP 2152
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2083 PVKLGGEAAHLPHLRPLPESQPSSSPLLQTAPGVKGHQRVVTLAQHISEVITQDYTRHHP 2142

Qy        2153 QQLSAPLPAPLYSFPGASCPVLDLRRPPSDLYLPPPDHGAPARGSPHSEGGKRSPEPNKT 2212
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2143 QQLSAPLPAPLYSFPGASCPVLDLRRPPSDLYLPPPDHGAPARGSPHSEGGKRSPEPNKT 2202

Qy        2213 SVLGGGEDGIEPVSPPEGMTEPGHSRSAVYPLLYRDGEQTEPSRMGSKSPGNTSQPPAFF 2272
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2203 SVLGGGEDGIEPVSPPEGMTEPGHSRSAVYPLLYRDGEQTEPSRMGSKSPGNTSQPPAFF 2262

Qy        2273 SKLTESNSAMVKSKKQEINKKLNTHNRNEPEYNISQPGTEIFNMPAITGTGLMTYRSQAV 2332
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2263 SKLTESNSAMVKSKKQEINKKLNTHNRNEPEYNISQPGTEIFNMPAITGTGLMTYRSQAV 2322

Qy        2333 QEHASTNMGLEAIIRKALMGKYDQWEESPPLSANAFNPLNASASLPAAMPITAADGRSDH 2392
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2323 QEHASTNMGLEAIIRKALMGKYDQWEESPPLSANAFNPLNASASLPAAMPITAADGRSDH 2382

Qy        2393 TLTSPGGGGKAKVSGRPSSRKAKSPAPGLASGDRPPSVSSVHSEGDCNRRTPLTNRVWED 2452
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2383 TLTSPGGGGKAKVSGRPSSRKAKSPAPGLASGDRPPSVSSVHSEGDCNRRTPLTNRVWED 2442

Qy        2453 RPSSAGSTPFPYNPLIMRLQAGVMASPPPPGLPAGSGPLAGPHHAWDEEPKPLLCSQYET 2512
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2443 RPSSAGSTPFPYNPLIMRLQAGVMASPPPPGLPAGSGPLAGPHHAWDEEPKPLLCSQYET 2502

Qy        2513 LSDSE 2517
               |||||
Db        2503 LSDSE 2507
```

RESULT 3
US-09-819-104A-5
; Sequence 5, Application US/09819104A
; Publication No. US20030027137A1
; GENERAL INFORMATION:
; APPLICANT: Chen, J. Don
; TITLE OF INVENTION: NOVEL NUCLEAR RECEPTOR COREPRESSOR MOLECULES
; TITLE OF INVENTION: AND USES THEREFOR
; FILE REFERENCE: UMG-030
; CURRENT APPLICATION NUMBER: US/09/819,104A
; CURRENT FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/193,138
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 2462
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-819-104A-5

Query Match 81.7%; Score 10796.5; DB 10; Length 2462;
Best Local Similarity 83.4%; Pred. No. 0;
Matches 2126; Conservative 94; Mismatches 210; Indels 119; Gaps 36;

```
Qy           1 MSGSTQLVAQTWRATEPRYPPHSLSYPVQIARTHTDVGLLEYQHHSRDYASHLSPGSIIQ 60
```





```
Qy       2190 HGAPARGSPHSEGGKRSPEPNKTSVLGGGEDGIEPVSPPEGMTEPGHSRSAVYPLLYRDG 2249
             || |||||||||||||||||:|||||   || ||||||||||||||:||||||||||||
Db       2135 HGTPARGSPHSEGGKRSPEPSKTSVLGSSEDAIEPVSPPEGMTEPGHARSAVYPLLYRDG 2194

Qy       2250 EQTEPSRMGSKSPGNTSQPPAFFSKLTESNSAMVKSKKQEINKKLNTHNRNEPEYNISQP 2309
             || || ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db       2195 EQGEP-RMGSKSPGNTSQPPAFFSKLTESNSAMVKSKKQEINKKLNTHNRNEPEYNIGQP 2253

Qy       2310 GTEIFNMPAITGTGLMTYRSQAVQEHASTNMGLEAIIRKALMGKYDQWEESPPLSANAFN 2369
             |||||||||||| |||| |||||||||||||||||||||||||||||||| ||| |||||
Db       2254 GTEIFNMPAITGAGLMTCRSQAVQEHASTNMGLEAIIRKALMGKYDQWEEPPPLGANAFN 2313

Qy       2370 PLNASASLP-AAMPITAADGRSDHTLTSPGGGGKAKVSGRPSSRKAKSPAPGLASGDRPP 2428
             ||||||||| |||||| ||||||| ||||||||||||||||||||||||||||||||||||
Db       2314 PLNASASLPAAAMPITTADGRSDHALTSPGGGGKAKVSGRPSSRKAKSPAPGLASGDRPP 2373

Qy       2429 SVSSVHSEGDCNRRTPLTNRVWEDRPSSAGSTPFPYNPLIMRLQAGVMASPPPPGLPAGS 2488
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db       2374 SVSSVHSEGDCNRRTPLTNRVWEDRPSSAGSTPFPYNPLIMRLQAGVMASPPPPGLAAGS 2433

Qy       2489 GPLAGPHHAWDEEPKPLLCSQYETLSDSE 2517
             |||||||||||||||||||||||||||||
Db       2434 GPLAGPHHAWDEEPKPLLCSQYETLSDSE 2462


RESULT 4
US-10-351-750-1
; Sequence 1, Application US/10351750
; Publication No. US20030138836A1
; GENERAL INFORMATION:
;  APPLICANT: THE SALK INSTITUTE FOR BIOLOGICAL STUDIES
;  APPLICANT:  EVANS, RONALD
;  APPLICANT:  CHEN, J.
;  TITLE OF INVENTION: TRANSCRIPTIONAL CO-REPRESSOR THAT INTERACTS WITH NUCLEAR HORMONE
;  TITLE OF INVENTION:  RECEPTORS
;  FILE REFERENCE: SALK1510-2
;  CURRENT APPLICATION NUMBER: US/10/351,750
;  CURRENT FILING DATE:  2003-01-23
;  PRIOR APPLICATION NUMBER: US/09/337,384
;  PRIOR FILING DATE: 1999-06-21
;  PRIOR APPLICATION NUMBER: 08/522,726
;  PRIOR FILING DATE: 1995-09-01
;  NUMBER OF SEQ ID NOS: 3
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
;   LENGTH: 1495
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-351-750-1

  Query Match             59.6%;  Score 7871;  DB 14;  Length 1495;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1487;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1031 DKEAFAAEAQKLPGDPPCWTSGLPFPVPPREVIKASPHAPDPSAFSYAPPGHPLPLGLHD 1090
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          9 DKEAFAAEAQKLPGDPPCWTSGLPFPVPPREVIKASPHAPDPSAFSYAPPGHPLPLGLHD 68

Qy       1091 TARPVLPRPPTISNPPPLISSAKHPSVLERQIGAISQGMSVQLHVPYSEHAKAPVGPVTM 1150
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         69 TARPVLPRPPTISNPPPLISSAKHPSVLERQIGAISQGMSVQLHVPYSEHAKAPVGPVTM 128

Qy       1151 GLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPESLGVPTAQEASVLRGTALGSVPGGSITK 1210
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        129 GLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPESLGVPTAQEASVLRGTALGSVPGGSITK 188

Qy       1211 GIPSTRVPSDSAITYRGSITHGTPADVLYKGTITRIIGEDSPSRLDRGREDSLPKGHVIY 1270
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        189 GIPSTRVPSDSAITYRGSITHGTPADVLYKGTITRIIGEDSPSRLDRGREDSLPKGHVIY 248

Qy       1271 EGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPHETAAPKRTYDMMEGRVGRAISSASIEGL 1330
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        249 EGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPHETAAPKRTYDMMEGRVGRAISSASIEGL 308

Qy       1331 MGRAIPPERHSPHHLKEQHHIRGSITQGIPRSYVEAQEDYLRREAKLLKREGTPPPPPPS 1390
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        309 MGRAIPPERHSPHHLKEQHHIRGSITQGIPRSYVEAQEDYLRREAKLLKREGTPPPPPPS 368

Qy       1391 RDLTEAYKTQALGPLKLKPAHEGLVATVKEAGRSIHEIPREELRHTPELPLAPRPLKEGS 1450
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        369 RDLTEAYKTQALGPLKLKPAHEGLVATVKEAGRSIHEIPREELRHTPELPLAPRPLKEGS 428

Qy       1451 ITQGTPLKYDTGASTTGSKKHDVRSLIGSPGRTFPPVHPLDVMADARALERACYEESLKS 1510
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        429 ITQGTPLKYDTGASTTGSKKHDVRSLIGSPGRTFPPVHPLDVMADARALERACYEESLKS 488

Qy       1511 RPGTASSSGGSIARGAPVIVPELGKPRQSPLTYEDHGAPFAGHLPRGSPVTMREPTPRLQ 1570
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        489 RPGTASSSGGSIARGAPVIVPELGKPRQSPLTYEDHGAPFAGHLPRGSPVTMREPTPRLQ 548

Qy       1571 EGSLSSSKASQDRKLTSTPREIAKSPHSTVPEHHPHPISPYEHLLRGVSGVDLYRSHIPL 1630
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        549 EGSLSSSKASQDRKLTSTPREIAKSPHSTVPEHHPHPISPYEHLLRGVSGVDLYRSHIPL 608

Qy       1631 AFDPTSIPRGIPLDAAAAYYLPRHLAPNPTYPHLYPPYLIRGYPDTAALENRQTIINDYI 1690
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        609 AFDPTSIPRGIPLDAAAAYYLPRHLAPNPTYPHLYPPYLIRGYPDTAALENRQTIINDYI 668

Qy       1691 TSQQMHHNTATAMAQRADMLRGLSPRESSLALNYAAGPRGIIDLSQVPHLPVLVPPTPGT 1750
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        669 TSQQMHHNTATAMAQRADMLRGLSPRESSLALNYAAGPRGIIDLSQVPHLPVLVPPTPGT 728

Qy       1751 PATAMDRLAYLPTAPQPFSSRHSSSPLSPGGPTHLTKPTTTSSSERERDRDRERDRDRER 1810
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        729 PATAMDRLAYLPTAPQPFSSRHSSSPLSPGGPTHLTKPTTTSSSERERDRDRERDRDRER 788

Qy       1811 EKSILTSTTTVEHAPIWRPGTEQSSGSSGSSGGGGGSSSRPASHSHAHQHSPISPRTQDA 1870
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        789 EKSILTSTTTVEHAPIWRPGTEQSSGSSGSSGGGGGSSSRPASHSHAHQHSPISPRTQDA 848

Qy       1871 LQQRPSVLHNTGMKGIITAVEPSKPTVLRSTSTSSPVRPAATFPPATHCPLGGTLDGVYP 1930
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        849 LQQRPSVLHNTGMKGIITAVEPSKPTVLRSTSTSSPVRPAATFPPATHCPLGGTLDGVYP 908

Qy       1931 TLMEPVLLPKEAPRVARPERPRADTGHAFLAKPPARSGLEPASSPSKGSEPRPLVPPVSG 1990
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        909 TLMEPVLLPKEAPRVARPERPRADTGHAFLAKPPARSGLEPASSPSKGSEPRPLVPPVSG 968

Qy       1991 HATIARTPAKNLAPHHASPDPPAPPASASDPHREKTQSKPFSIQELELRSLGYHGSSYSP 2050
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        969 HATIARTPAKNLAPHHASPDPPAPPASASDPHREKTQSKPFSIQELELRSLGYHGSSYSP 1028

Qy       2051 EGVEPVSPVSSPSLTHDKGLPKHLEELDKSHLEGELRPKQPGPVKLGGEAAHLPHLRPLP 2110
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1029 EGVEPVSPVSSPSLTHDKGLPKHLEELDKSHLEGELRPKQPGPVKLGGEAAHLPHLRPLP 1088

Qy       2111 ESQPSSSPLLQTAPGVKGHQRVVTLAQHISEVITQDYTRHHPQQLSAPLPAPLYSFPGAS 2170
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1089 ESQPSSSPLLQTAPGVKGHQRVVTLAQHISEVITQDYTRHHPQQLSAPLPAPLYSFPGAS 1148

Qy       2171 CPVLDLRRPPSDLYLPPPDHGAPARGSPHSEGGKRSPEPNKTSVLGGGEDGIEPVSPPEG 2230
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1149 CPVLDLRRPPSDLYLPPPDHGAPARGSPHSEGGKRSPEPNKTSVLGGGEDGIEPVSPPEG 1208

Qy       2231 MTEPGHSRSAVYPLLYRDGEQTEPSRMGSKSPGNTSQPPAFFSKLTESNSAMVKSKKQEI 2290
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1209 MTEPGHSRSAVYPLLYRDGEQTEPSRMGSKSPGNTSQPPAFFSKLTESNSAMVKSKKQEI 1268

Qy       2291 NKKLNTHNRNEPEYNISQPGTEIFNMPAITGTGLMTYRSQAVQEHASTNMGLEAIIRKAL 2350
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1269 NKKLNTHNRNEPEYNISQPGTEIFNMPAITGTGLMTYRSQAVQEHASTNMGLEAIIRKAL 1328

Qy       2351 MGKYDQWEESPPLSANAFNPLNASASLPAAMPITAADGRSDHTLTSPGGGGKAKVSGRPS 2410
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1329 MGKYDQWEESPPLSANAFNPLNASASLPAAMPITAADGRSDHTLTSPGGGGKAKVSGRPS 1388
```

```
Qy       2411 SRKAKSPAPGLASGDRPPSVSSVHSEGDCNRRTPLTNRVWEDRPSSAGSTPFPYNPLIMR 2470
Db       1389 SRKAKSPAPGLASGDRPPSVSSVHSEGDCNRRTPLTNRVWEDRPSSAGSTPFPYNPLIMR 1448

Qy       2471 LQAGVMASPPPPGLPAGSGPLAGPHHAWDEEPKPLLCSQYETLSDSE 2517
Db       1449 LQAGVMASPPPPGLPAGSGPLAGPHHAWDEEPKPLLCSQYETLSDSE 1495


RESULT 5
US-10-087-192-651
; Sequence 651, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
;  APPLICANT: Morris, David W.
;  APPLICANT:  Engelhard, Eric K.
;  TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
;  TITLE OF INVENTION:  CANCER
;  FILE REFERENCE: 529452000122
;  CURRENT APPLICATION NUMBER: US/10/087,192
;  CURRENT FILING DATE:  2002-03-01
;  PRIOR APPLICATION NUMBER: US 09/747,377
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: US 09/798,586
;  PRIOR FILING DATE: 2001-03-02
;  NUMBER OF SEQ ID NOS: 2059
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 651
;   LENGTH: 1585
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-087-192-651

  Query Match             50.6%;  Score 6685;  DB 12;  Length 1585;
  Best Local Similarity   79.4%;  Pred. No. 2.4e-298;
  Matches 1327;  Conservative   72;  Mismatches  155;  Indels  118;  Gaps   26;

Qy        295 KQKFCQRYDQLMEALEKKVERIENNPRRRAKESKVREYYEKQFPEIRKQRELQERMQSRV 354
Db          1 EQRFCQRYDQLMEAWEKKVERIENNPRRRAKESKVREYYEKQFPEIRKQRELQERMQSRV 60

Qy        355 GQRGSGLSMSAARSEHEVSEIIDGLSEQENLEKQMRQLAVIPPMLYDADQQRIKFINMNG 414
Db         61 GQRGSGLSMSAARSEHEVSEIIDGLSEQENLEKQMRQLAVIPPMLYDADQQRIKFINMNG 120

Qy        415 LMADPMKVYKDRQVMNMWSEQEKETFREKFMQHPKNFGLIASFLERKTVAECVLYYYLTK 474
Db        121 LMDDPMKVYKDRQVTNMWSEQERDTFREKFMQHPKNFGLIASFLERKTVAECVLYYYLTK 180

Qy        475 KNENYKSLVRRSYRRRGKSQQQQQQQQQQQQQQQQQPMPRSSQEEKDEKEKEKEAEKEEE 534
Db        181 KNENYKSLVRRSYRRRGKSQQQQQQQQQQQQQQ----MARSSQEEKEEKEKEKEADKEEE 236

Qy        535 KPEVENDKEDLLKEKTDDTSGEDNDEKEAVASKGRKTANSQGRRKGRITRSMANEANSEE 594
Db        237 KQDAENEKEELSKEKTDDTSGEDNDEKEAVASKGRKTANSQGRRKGRITRSMANEANHEE 296

Qy        595 AITPQQSAELASMELNESSRWTEEEMETAKKGLLEHGRNWSAIARMVGSKTVSQCKNFYF 654
Db        297 TATPQQSSELASMEMNESSRWTEEEMETAKKGLLEHGRNWSAIARMVGSKTVSQCKNFYF 356

Qy        655 NYKKRQNLDEILQQHKLKMEKERNARRKKKKAPAAASEEAAFPPVVEDEEMEASGVSGNE 714
Db        357 NYKKRQNLDEILQQHKLKMEKERNARRKKKKTPAAASEETAFPPAAEDEEMEASGASANE 416

Qy        715 EEMVEEAEALHASGNEVPR-GECSGPATVNNSSDTESIPSPHTEAAKDTGQNGPKPPAT- 772
Db        417 EELAEEAEASQASGNEVPRVGECSGPAAVNNSSDTESVPSPRSEATKDT---GPKPTGTE 473

Qy        773 -LGADGPPPGPPTPPRRTSRAPIEPTPASEATGAPTPPPAPPSPSAPPPVVPKEEKEEET 831
Db        474 ALPAATQPPVP--PPEEPAAAPAEPSPVPDASGPPSPEPS-PSPAAPPATVDKDEQEAPA 530

Qy        832 AAAPPVEEGEEQKPPAAEELAVDTGKAEEP---------VKSECTEEAEEGPA-KGKDAE 881
Db        531 APAPQTEDAKEQKSEAEE---IDVGKPEEPEASEEPPESVKSDHKEETEEEPEDKAKGTE 587

Qy        882 AAEATAEGALKAEKKEGGSGRATT-AKSSGAPQDSDSSATCSADEVDEAEGGDKNRLLSP 940
Db        588 AIETVSEAPLKVE--EAGSKAAVTKGSSSGATQDSDSSATCSADEVDEPEGGDKGRLLSP 645

Qy        941 RPSLLTPTGDPRANASPQKPLDLKQLKQRAAAIPPIQVTKVHEPPREDAAPTKPAPPAPP 1000
Db        646 RPSLLTPAGDPRASTSPQKPLDLKQLKQRAAAIPPI-VTKVHEPPREDTVPPKPVPPVPP 704

Qy       1001 PPQNLQPESDAPQQPGSSPRGKSRSPAPPADKE--------AFAAEAQKLPGDPPCWTSG 1052
Db        705 PTQHLQPEGDVSQQSGGSPRGKSRSPVPPAEKEAEKPAFFPAFPTEGPKLPTEPPRWSSG 764

Qy       1053 LPFPVPPREVIKASPHAPDPSAFSYAPPGHPLPLGLHDTARPVLPRPPTISNPPPLISSA 1112
Db        765 LPFPIPPREVIKTSPHAADPSAFSYTPPGHPLPLGLHDSARPVLPRPP-ISNPPPLISSA 823

Qy       1113 KHPSVLERQIGAISQGMSVQLHVPYSEHAKAPVGPVTMGLPLPMDPKKLAPFSGVKQEQL 1172
Db        824 KHPGVLERQLGAISQGMSVQLRVPHSEHAKAPMGPLTMGLPLAVDPKKL----------- 872

Qy       1173 SPRGQAGPPESLGVPTAQEASVLRGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHG 1232
Db        873 ----------------------GTALGSATSGSITKGLPSTRAADGP--SYRGSITHG 906

Qy       1233 TPADVLYKGTITRIIGEDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKE 1292
Db        907 TPADVLYKGTISRIVGEDSPSRLDRAREDTLPKGHVIYEGKKGHVLSYEGGMSVSQCSKE 966

Qy       1293 DGRSSSGPPHETAAPKRTYDMMEGRVGRAISSASIEGLMGRAIPPERHSPHHLKEQHHIR 1352
Db        967 DGRSSSGPPHETAAPKRTYDMMEGRVGRTVTSASIEGLMGRAI-PEQHSP-HLKEQHHIR 1024

Qy       1353 GSITQGIPRSYVEAQEDYLRREAKLLKREGTPPPPPPSRDLTEAYKTQ---ALGPLKLKP 1409
Db       1025 GSITQGIPRSYVEAQEDYLRREAKLLKREGTPPPPPPRDLTETYKPRPLDPLGPLKLKP 1084

Qy       1410 AHEGLVATVKEAGRSIHEIPREELRHTPELPLAPRPLKEGSITQGTPLKYDTGASTTGSK 1469
Db       1085 THEGVVATVKEAGRSIHEIPREELRRTPELPLAPRPLKEGSITQGTPLKYDSGAPSTGTK 1144

Qy       1470 KHDVRSLIGSPGRTFPPVHPLDVMADARALERACYEESLKSRPGTASSSGGSIARGAPVI 1529
Db       1145 KHDVRSIIGSPGRPFPALHPLDIMADARALERACYEESLKSRSGTSSGAGGSITRGAPVV 1204

Qy       1530 VPELGKPRQSPLTYEDHGAPFAGHLPRGSPVTMREPTPRLQEGSLSSSKASQDRKLTSTP 1589
Db       1205 VPELGKPRQSPLTYEDHGAPFTSHLPRGSPVTTREPTPRLQEGSLLSSKASQDRKLTSTP 1264

Qy       1590 REIAKSPHSTVPEHHPHPISPYEHLLRGVSGVDLYRSHIPLAFDPTSIPRGIPLD-AAAA 1648
Db       1265 REIAKSPHSTVPEHHPHPISPYEHLLRGVTGVDLYRGHIPLAFDPTSIPRGIPLEAAAAA 1324

Qy       1649 YYLPRHLAPNPTYPHLYPPYLIRGYPDTAALENRQTIINDYITSQQMHHNTATAMAQRAD 1708
Db       1325 YYLPRHLAPSPTYPHLYPPYLIRGYPDTAALENRQTIINDYITSQQMHHNAASAMAQRAD 1384

Qy       1709 MLRGLSPRESSLALNYAAGPRGIIDLSQVPHLPVLVPPTPGTPATAMDRLAYLPTAPQPF 1768
Db       1385 MLRGLSPRESSLALNYAAGPRGIIDLSQVPHLPVLVPPTPGTPATAIDRLAYLPTAPPPF 1444

Qy       1769 SSRHSSSPLSPGGPTHLTKPTTTSSSERERDRDRERDRDREREKSILTSTTTVEHAPIWR 1828
Db       1445 SSRHSSSPLSPGGPTHLAKPTATSSSERERERERERD------KSILTSTTTVEHAPIWR 1498

Qy       1829 PGTEQSSGSSGSSGGGGSSSRPASHSHAHQHSPISPRTQDALQQRPSVLHNTGMKGIIT 1888
Db       1499 PGTEQSSGA--------GGSSRPASHT--HQHSPISPRTQDALQQRPSVLHNTSMKGVVT 1548
```

```
Qy       1889 AVEPSKPTVLRSTSTSSPVRPAA-----TFPPATHCPLGGTLDGVYPTLMEP 1935
                                          ||: | |    |    ||   |
Db       1549 ----------SRGTRHAHGPEVHLHLFACPPSCHIPTCHPL----PTWWHP 1585
```

RESULT 6
US-10-146-473-54
; Sequence 54, Application US/10146473
; Publication No. US20030108888A1
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew
; APPLICANT: Gout, Ivan
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Gure, Ali
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Old, Lloyd
; TITLE OF INVENTION: Breast Cancer Antigens
; FILE REFERENCE: L00461/70130(JRV)
; CURRENT APPLICATION NUMBER: US/10/146,473
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: US 60/291,150
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 54
; LENGTH: 876
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-146-473-54

Query Match 34.9%; Score 4611; DB 14; Length 876;
Best Local Similarity 94.7%; Pred. No. 1.7e-203;
Matches 873; Conservative 0; Mismatches 3; Indels 46; Gaps 1;

```
Qy       1596 PHSTVPEHHPHPISPYEHLLRGVSGVDLYRSHIPLAFDPTSIPRGIPLDAAAAYYLPRHL 1655
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 PHSTVPEHHPHPISPYEHLLRGVSGVDLYRSHIPLAFDPTSIPRGIPLDAAAAYYLPRHL 60

Qy       1656 APNPTYPHLYPPYLIRGYPDTAALENRQTIINDYITSQQMHHNTATAMAQRADMLRGLSP 1715
              |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db         61 APNPTYPHLYPPYLIRGYPDTAALENRQTIINDYITSQQMHHNAATAMAQRADMLRGLSP 120

Qy       1716 RESSLALNYAAGPRGIIDLSQVPHLPVLVPPTPGTPATAMDRLAYLPTAPQPFSSRHSSS 1775
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RESSLALNYAAGPRGIIDLSQVPHLPVLVPPTPGTPATAMDRLAYLPTAPQPFSSRHSSS 180

Qy       1776 PLSPGGPTHLTKPTTTSSSERERDRDRERDRDRDREREKSILTSTTTVEHAPIWRPGTEQSS 1835
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PLSPGGPTHLTKPTTTSSSERERDRDRERDRDRDREREKSILTSTTTVEHAPIWRPGTEQSS 240

Qy       1836 GSSGSSGGGGGSSSRPASHSHAHQHSPISPRTQDALQQRPSVLHNTGMKGIITAVEPSKP 1895
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        241 GSSGSSGGGGGSSSRPASHSHAHQHSPISPRTQDALQQRPSVLHNTGMKGIITAVEPSTP 300

Qy       1896 TVLRSTSTSSPVRPAATFPPATHCPLGGTLDGVYPTLMEPVLLPKEAPRVARPERPRADT 1955
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TVLRSTSTSSPVRPAATFPPATHCPLGGTLDGVYPTLMEPVLLPKEAPRVARPERPRADT 360

Qy       1956 GHAFLAKPPARSGLEPASSPSKGSEPRPLVPPVSGHATIARTPAKNLAPHHASPDPPAPP 2015
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GHAFLAKPPARSGLEPASSPSKGSEPRPLVPPVSGHATIARTPAKNLAPHHASPDPPAPP 420

Qy       2016 ASASDPHREKTQSKPFSIQELELRSLGYHGSSYSPEGVEPVSPVSSPSLTHDKGLPKHLE 2075
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ASASDPHREKTQSKPFSIQELELRSLGYHGSSYSPEGVEPVSPVSSPSLTHDKGLPKHLE 480

Qy       2076 ELDKSHLEGELRPKQPGPVKLGGEAAHLPHLRPLPESQPSSSPLLQTAPGVKGHQRVVTL 2135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ELDKSHLEGELRPKQPGPVKLGGEAAHLPHLRPLPESQPSSSPLLQTAPGVKGHQRVVTL 540

Qy       2136 AQHISEVITQDYTRHHPQQLSAPLPAPLYSFPGASCPVLDLRRPPSDLYLPPPDHGAPAR 2195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AQHISEVITQDYTRHHPQQLSAPLPAPLYSFPGASCPVLDLRRPPSDLYLPPPDHGAPAR 600

```
Qy       2196 GSPHSEGGKRSPEPNKTSVLGGGEDGIEPVSPPEGMTEPGHSRSAVYPLLYRDGEQTEPS 2255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GSPHSEGGKRSPEPNKTSVLGGGEDGIEPVSPPEGMTEPGHSRSAVYPLLYRDGEQTEPS 660

Qy       2256 RMGSKSPGNTSQPPAFFSKLTESNSAMVKSKKQEINKKLNTHNRNEPEYNISQPGTEIFN 2315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 RMGSKSPGNTSQPPAFFSKLTESNSAMVKSKKQEINKKLNTHNRNEPEYNISQPGTEIFN 720

Qy       2316 MPAITGTGLMTYRSQAVQEHASTNMGLEAIIRKALMGKYDQWEESPPLSANAFNPLNASA 2375
              ||||||||||||||||||||||||||||||||||||||
Db        721 MPAITGTGLMTYRSQAVQEHASTNMGLEAIIRKALM---------------------- 756

Qy       2376 SLPAAMPITAADGRSDHTLTSPGGGGKAKVSGRPSSRKAKSPAPGLASGDRPPSVSSVHS 2435
                                    ||||||||||||||||||||||||||||||||||||||
Db        757 ----------------------GGGGKAKVSGRPSSRKAKSPAPGLASGDRPPSVSSVHS 794

Qy       2436 EGDCNRRTPLTNRVWEDRPSSAGSTPFPYNPLIMRLQAGVMASPPPPGLPAGSGPLAGPH 2495
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        795 EGDCNRRTPLTNRVWEDRPSSAGSTPFPYNPLIMRLQAGVMASPPPPGLPAGSGPLAGAH 854

Qy       2496 HAWDEEPKPLLCSQYETLSDSE 2517
              |||||||||||||||||||||
Db        855 HAWDEEPKPLLCSQYETLSDSE 876
```

RESULT 7
US-10-341-434-236
; Sequence 236, Application US/10341434
; Publication No. US20030215835A1
; GENERAL INFORMATION:
; APPLICANT: OriGene Technologies
; TITLE OF INVENTION: Differentially Regulated Prostate Cancer Genes
; FILE REFERENCE: 9U 204 205 R1
; CURRENT APPLICATION NUMBER: US/10/341,434
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/348,164
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: US 60/348,119
; PRIOR FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 236
; LENGTH: 2440
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-341-434-236

Query Match 31.5%; Score 4165.5; DB 15; Length 2440;
Best Local Similarity 40.8%; Pred. No. 1.6e-182;
Matches 1103; Conservative 344; Mismatches 777; Indels 477; Gaps 109;

```
Qy         16 EPRYPPHSLSYPVQIARTHTDVGLLEYQ--HHSRDYASHLSPGSIIQPQRRRPSLLSEFQ 73
              : |||||: |     |   :  : :|:  |     || |       |  ||||||||||
Db         17 QSRYPPHSVQYTFPNTRHQQEFAVPDYRSSHLEVSQASQLLQQQQQQQLRRRPSLLSEFH 76

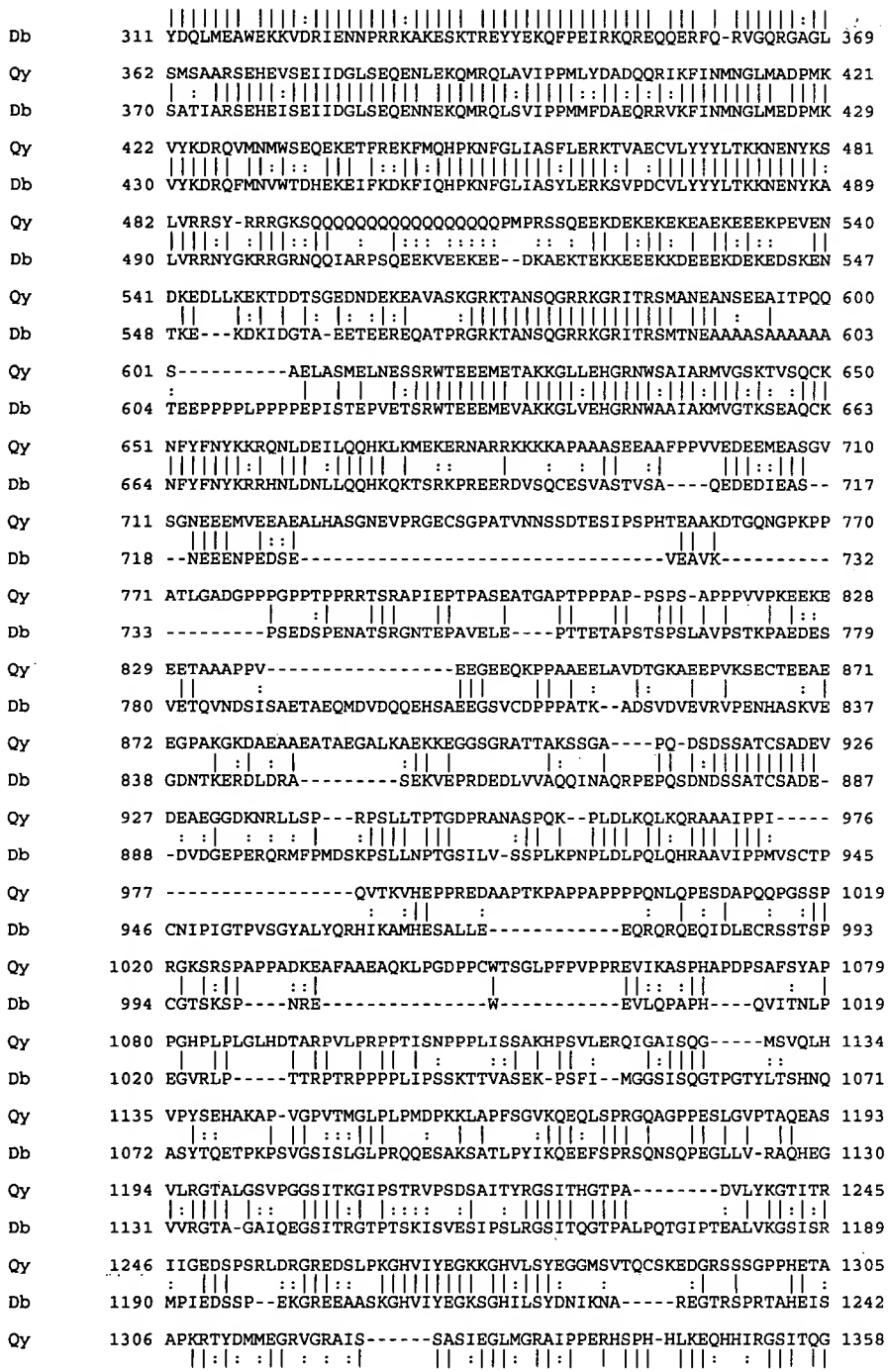
Qy         74 PGNERSQELHLRPESHSYLPELGKSEMEFIESKRPRLELLPD----------PLLRPSP 122
              ||:: | ||   |     : |     : : :|||||||| : |          ||: | |
Db         77 PGSDRPQE--RRTSYEPFHPGPSPVDHDSLESKRPRLEQVSDSHFQRVSAAVLPLVHPLP 134

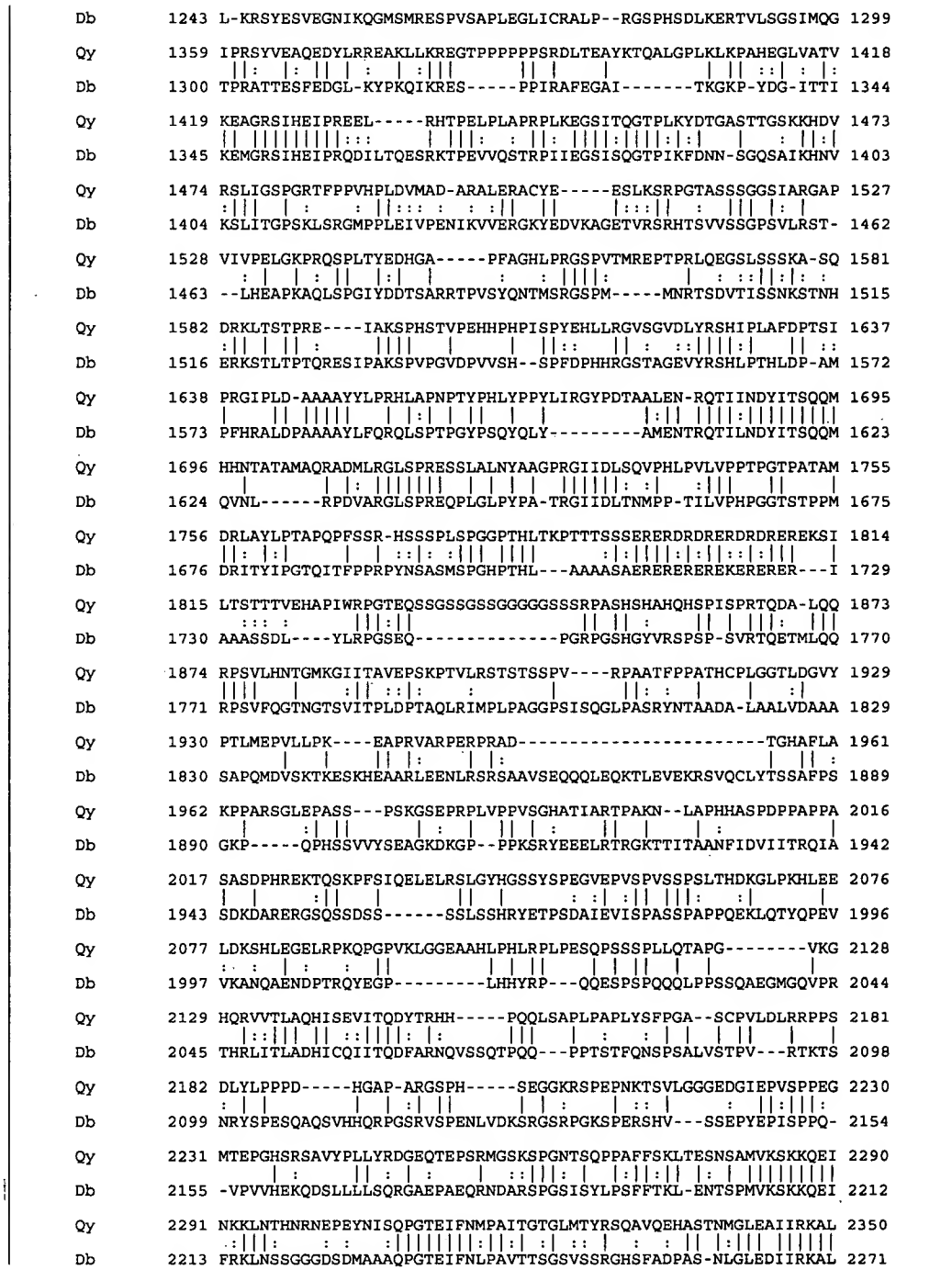
Qy        123 LLATGQPAGSEDLTKDRSLTGKLE-PVSPPSPPHTDPELELVPPRLSKEELIQNMDRVDR 181
                 |  | |  || :  || | | || |      :    |:|||||||:||||||
Db        135 ---EGLRA-SADAKKDPAFGGKHEAPSSPISGQPCGDDQNASPSKLSKEELIQSMDRVDR 190

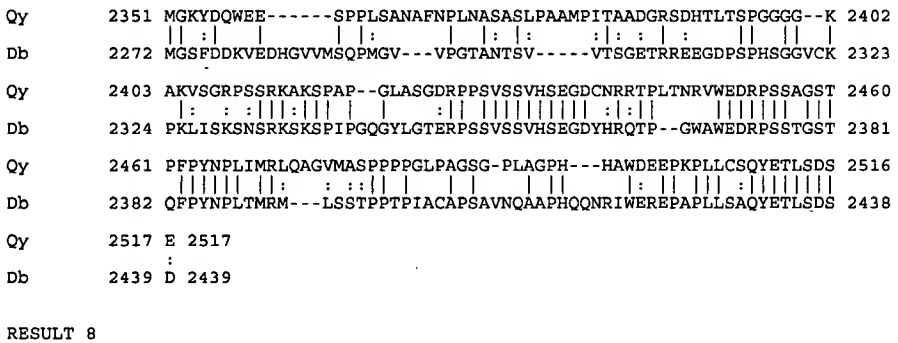
Qy        182 EITMVEQQISKLKKKQQQLEEEAAKPPEPEKPVSPPPIESKHRSLVQIIYDENRKKAEAA 241
              ||  ||||| |||||||||||||||||||||||:| ||||:|||||||||||| |
Db        191 EIAKVEQQILKLKKKQQQLEEEAAKPPEPEKPVSPPPVEQKHRSIVQIIYDENRKKAEEA 250

Qy        242 HRILEGLGPQVELPLYNQPSDTRQYHENIKINQAMRKKLILYFKRRNHARKQWKQKFCQR 301
              |:| |||||:|||||||||||: |||||| || |||||||:|||||||||| :|| |||
Db        251 HKIFEGLGPKVELPLYNQPSDTKVYHENIKTNQVMRKKLILFFKRRNHARKQREQKICQR 310

Qy        302 YDQLMEALEKKVERIENNPRRRAKESKVREYYEKQFPEIRKQRELQERMQSRVGQRGSGL 361
```
```

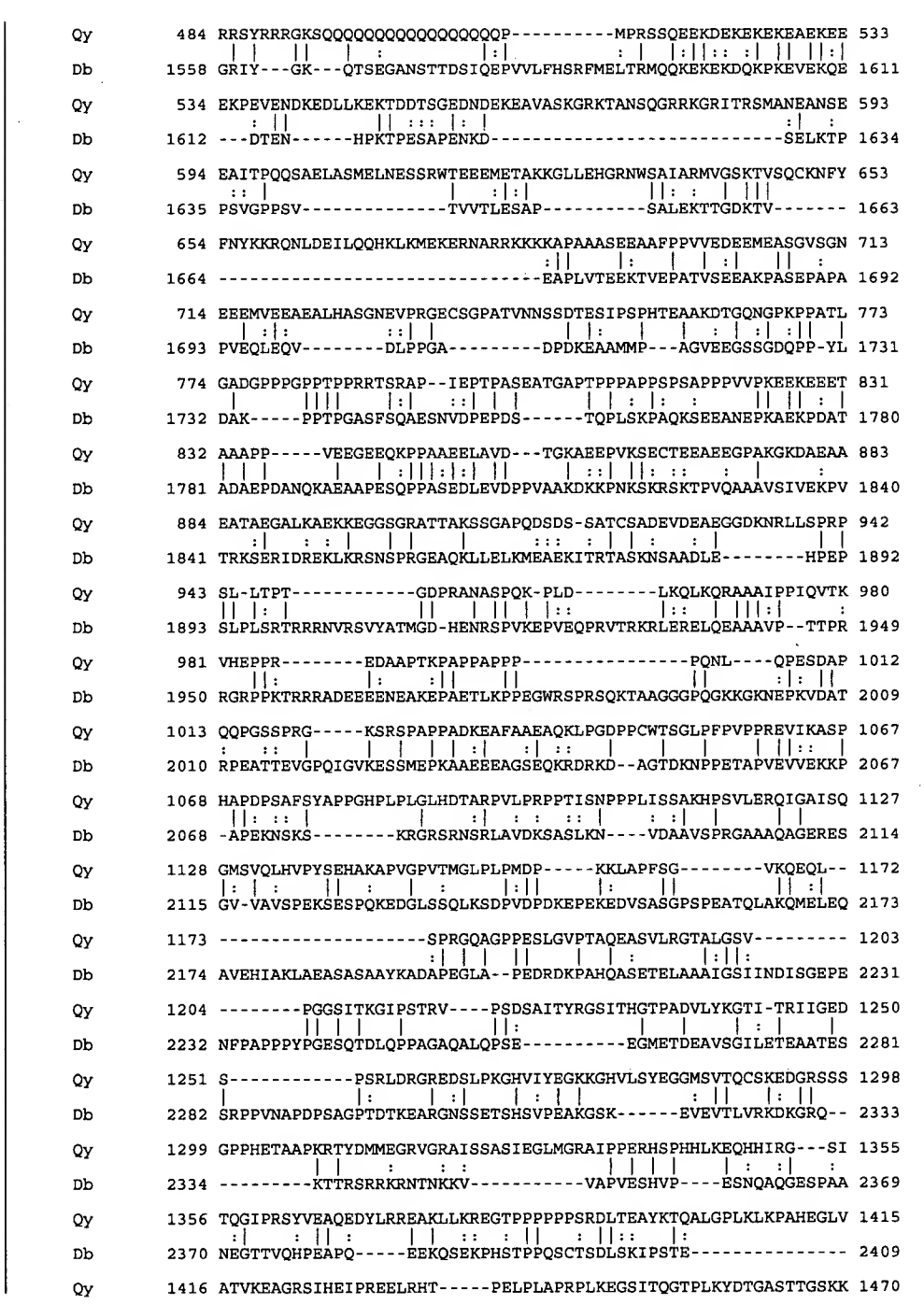






RESULT 8
US-10-263-929-143
; Sequence 143, Application US/10263929
; Publication No. US20040067535A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jaeseob
; APPLICANT: Galant, Ron
; TITLE OF INVENTION: Alzheimer's Disease Linked Genes
; FILE REFERENCE: LSD-07417
; CURRENT APPLICATION NUMBER: US/10/263,929
; CURRENT FILING DATE: 2002-10-03
; NUMBER OF SEQ ID NOS: 213
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 143
; LENGTH: 3664
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-263-929-143

Query Match 4.3%; Score 569.5; DB 12; Length 3664;
Best Local Similarity 19.6%; Pred. No. 2.5e-17;
Matches 568; Conservative 349; Mismatches 973; Indels 1011; Gaps 137;

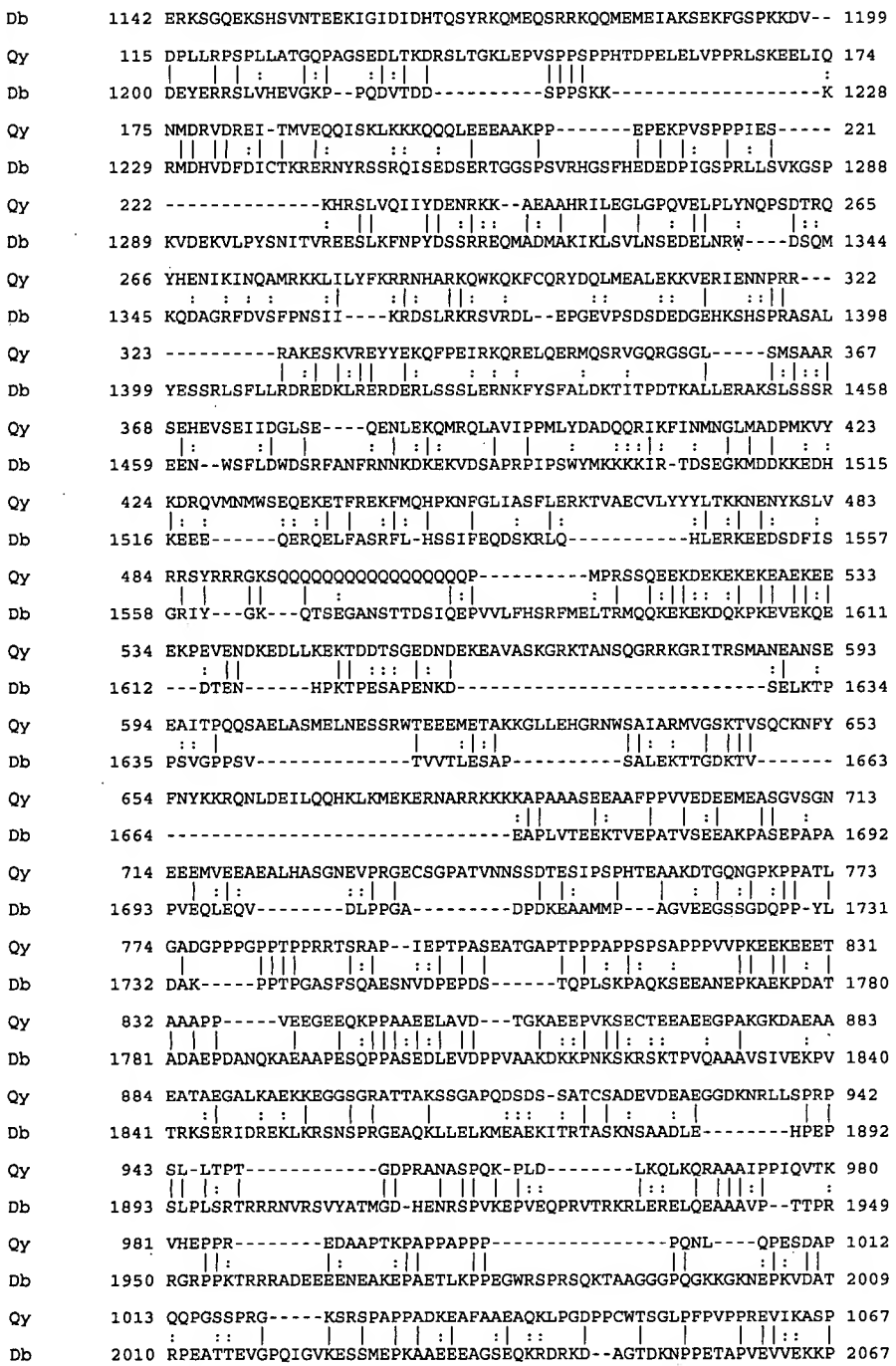


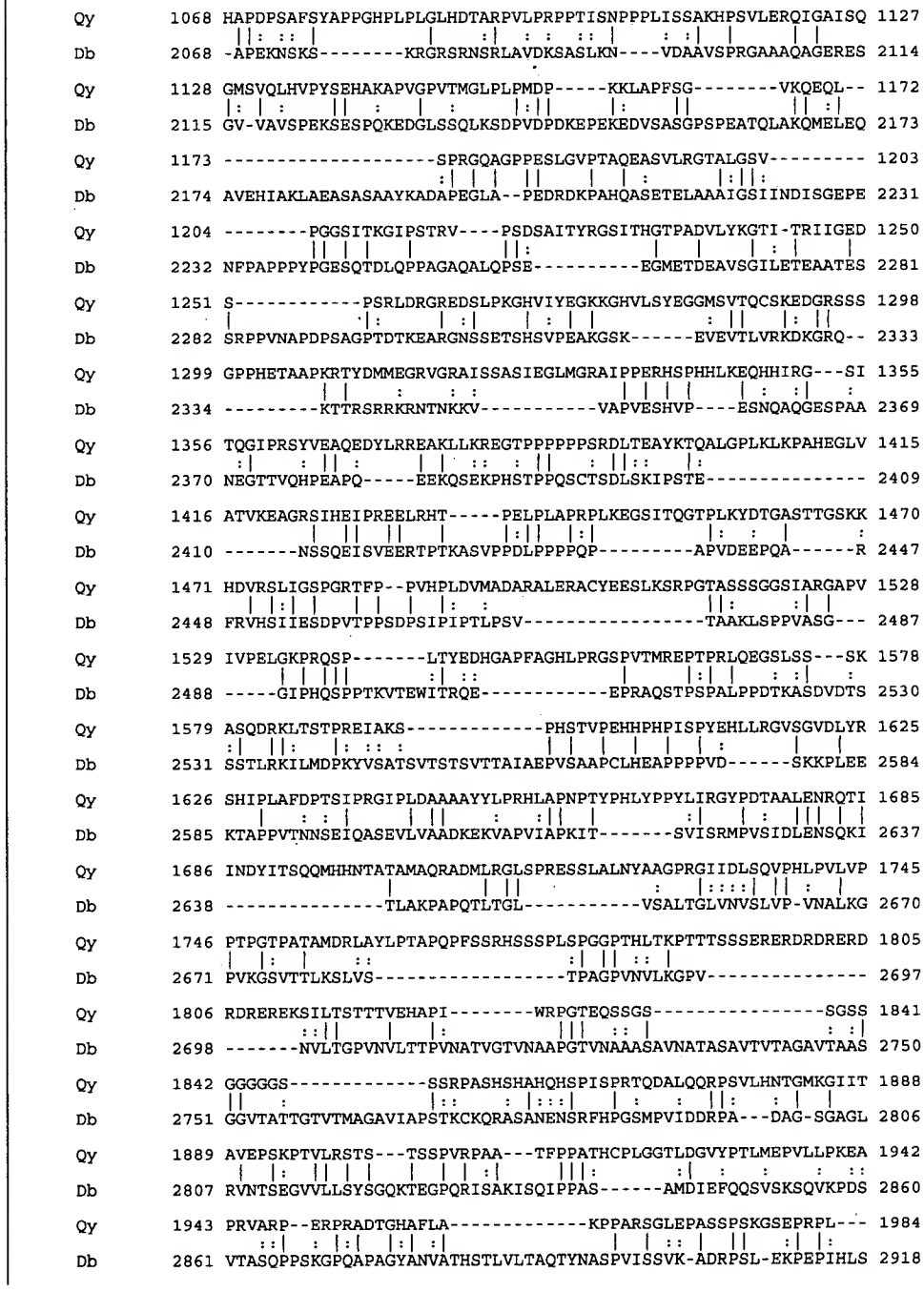
```
Db       2410 -------NSSQEISVEERTPTKASVPPDLPPPPQP---------APVDEEPQA------R 2447

Qy       1471 HDVRSLIGSPGRTFP--PVHPLDVMADARALERACYEESLKSRPGTASSSGGSIARGAPV 1528

Db       2448 FRVHSIIESDPVTPPSDPSIPIPTLPSV-----------------TAAKLSPPVASG--- 2487

Qy       1529 IVPELGKPRQSP-------LTYEDHGAPFAGHLPRGSPVTMREPTPRLQEGSLSS---SK 1578

Db       2488 -----GIPHQSPPTKVTEWITRQE------------EPRAQSTPSPALPPDTKASDVDTS 2530

Qy       1579 ASQDRKLTSTPREIAKS-------------PHSTVPEHHPHPISPYEHLLRGVSGVDLYR 1625

Db       2531 SSTLRKILMDPKYVSATSVTSTSVTTAIAEPVSAAPCLHEAPPPPVD------SKKPLEE 2584

Qy       1626 SHIPLAFDPTSIPRGIPLDAAAAYYLPRHLAPNPTYPHLYPPYLIRGYPDTAALENRQTI 1685

Db       2585 KTAPPVTNNSEIQASEVLVAADKEKVAPVIAPKIT-------SVISRMPVSIDLENSQKI 2637

Qy       1686 INDYITSQQMHHNTATAMAQRADMLRGLSPRESSLALNYAAGPRGIIDLSQVPHLPVLVP 1745

Db       2638 ---------------TLAKPAPQTLTGL-----------VSALTGLVNVSLVP-VNALKG 2670

Qy       1746 PTPGTPATAMDRLAYLPTAPQPFSSRHSSSPLSPGGPTHLTKPTTTSSSERERDRDRERD 1805

Db       2671 PVKGSVTTLKSLVS-----------------TPAGPVNVLKGPV--------------- 2697

Qy       1806 RDREREKSILTSTTTVEHAPI--------WRPGTEQSSGS----------------SGSS 1841

Db       2698 -------NVLTGPVNVLTTPVNATVGTVNAAPGTVNAAASAVNATASAVTVTAGAVTAAS 2750

Qy       1842 GGGGGS-------------SSRPASHSHAHQHSPISPRTQDALQQRPSVLHNTGMKGIIT 1888

Db       2751 GGVTATTGTVTMAGAVIAPSTKCKQRASANENSRFHPGSMPVIDDRPA---DAG-SGAGL 2806

Qy       1889 AVEPSKPTVLRSTS---TSSPVRPAA---TFPPATHCPLGGTLDGVYPTLMEPVLLPKEA 1942

Db       2807 RVNTSEGVVLLSYSGQKTEGPQRISAKISQIPPAS------AMDIEFQQSVSKSQVKPDS 2860

Qy       1943 PRVARP--ERPRADTGHAFLA-------------KPPARSGLEPASSPSKGSEPRPL--- 1984

Db       2861 VTASQPPSKGPQAPAGYANVATHSTLVLTAQTYNASPVISSVK-ADRPSL-EKPEPIHLS 2918

Qy       1985 -------------------VPPVSGHATIARTPA------------------------- 1999

Db       2919 VSTPVTQGGTVKVLTQGINTPPVLVHNQLVLTPSIVTTNKKLADPVTLKIETKVLQPANL 2978

Qy       2000 -KNLAPHHASPDPPAPPASASDP--HREKTQSKPF--SIQELELRSLGYHGSSYSPEGVE 2054

Db       2979 GSTLTPHH----PPALPSKLPTEVNHVPSGPSIPADRTVSHLAAAKLDAHSPRPSGPGPS 3034

Qy       2055 PVSPVSSPSLTHDKGL----------------------------PKH--LEELDKSHL- 2082

Db       3035 SFPRASHPSSTASTALSTNATVMLAAGIPVPQFISSIHPEQSVIMPPHSITQTVSLSHLS 3094

Qy       2083 EGELRPKQPG----PVKLGGEAAHLPHLRPLPESQPSSSPLLQTAPGVKGHQRVVTLAQH 2138

Db       3095 QGEVRMNTPTLPSITYSIRPEALHSPR-APL---QP---------------QQIEVRA-- 3133

Qy       2139 ISEVITQDYTRHHPQQLSAPLPAPLYSFPGASCPVLDLRRPPSD---LYLPPPDHGAPAR 2195

Db       3134 -------------PQRASTPQPAP------AGVPALASQHPPEEEVHYHLPVARATAPVQ 3174

Qy       2196 GS------------------------PH----------SEGGKRSPEPNKTSVLGGGEDG 2221

Db       3175 SEVLVMQSEYRLHPYTVPRDVRIMVHPHVTAVSEQPRAADGVVKVPPASKAP----QQPG 3230

Qy       2222 IEPVSPPEGMTEPGHSRSAV-------YPLLYRDGEQTEPSRMGSKSPGNTSQ-----PP 2269

Db       3231 KEAAKTPDAKAAPTPTPAPVPVPVPLPAPAPAPHGE----ARILTVTPSNQLQGLPLTPP 3286

Qy       2270 AFFSKLTESNSAMVKSKKQEINKKLNTHNRNE--PEYNISQPGTEIFNMPAITGTGLMTY 2327

Db       3287 ------------VVVTHGVQI-----VHSSGELFQEYRYGD---------------IRTY 3314

Qy       2328 RSQAVQEH----ASTNMGLEAIIRKALMGKYDQWEE-SPPLSANAFNPLNASASLPAAMP 2382

Db       3315 HPPAQLTHTQFPAASSVGLPSRTKTAAQGPPPEGEPLQPPQPVQSTQPAQPAPPCPPSQ- 3373

Qy       2383 ITAADGRSDHTLTSPGGGGKAK---VSGRPSSRKAKSPAPGLASG--DRPP----SVSSV 2433

Db       3374 -----------LGQPGQPPSSKMPQVSQEAKGTQTGVEQPRLPAGPANRPPEPHTQVQRA 3422

Qy       2434 HSE-GDCNRRTPLTNRVWEDRPSSAGSTPFPYNPLIMRLQAGVMASPPPPGLPAGSGPLA 2492

Db       3423 QAETGPTSFPSPVSVSMKPDLPVSLPTQTAPKQPLFVPTTSG-------PSTPPG---LV 3472

Qy       2493 GPHHAWDEEPK----PLLCSQ 2509

Db       3473 LPHTEFQPAPKQDSSPHLTSQ 3493
```

RESULT 9
US-10-177-293-423
; Sequence 423, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Gannavarpu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Pusztai, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
; TITLE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 423
; LENGTH: 3664
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-293-423

Query Match 4.3%; Score 569.5; DB 14; Length 3664;
Best Local Similarity 19.6%; Pred. No. 2.5e-17;
Matches 568; Conservative 349; Mismatches 973; Indels 1011; Gaps 137;

```
Qy         71 EFQPGNERSQELHLRPE-------SHSYLPELGKS------EMEFIESKR---PRLELLP 114
```



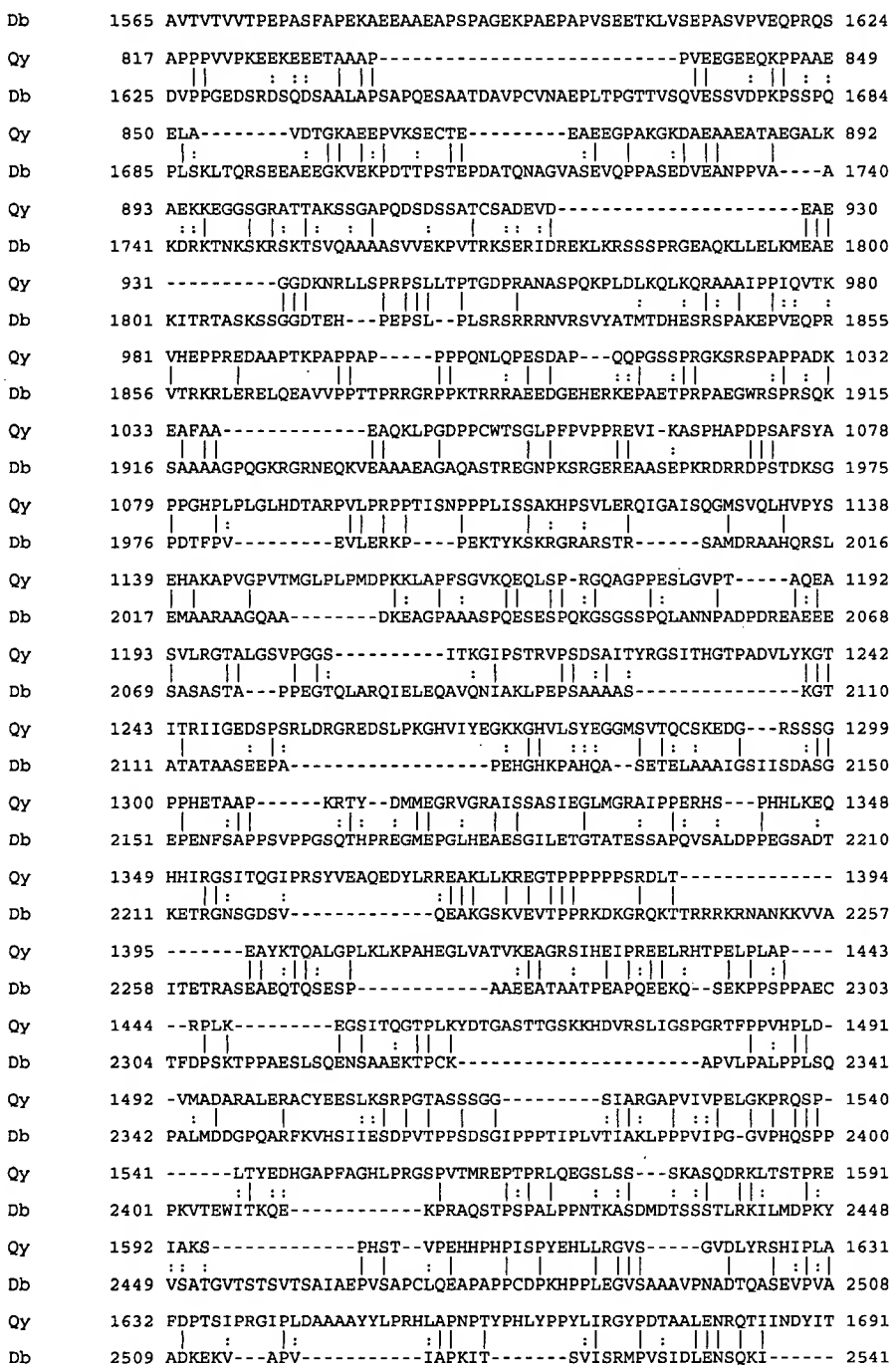


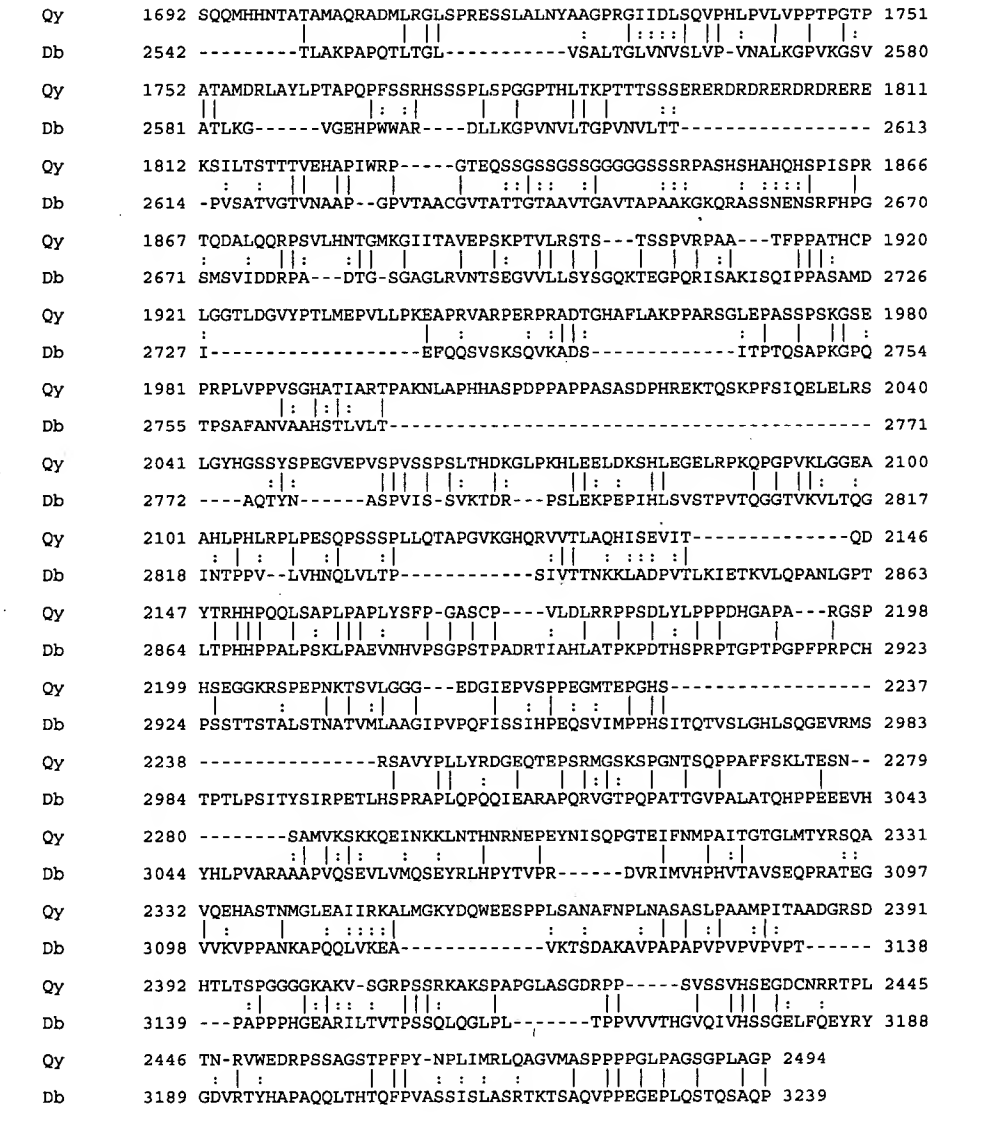
```
Qy       1985 ------------------VPPVSGHATIARTPA------------------------ 1999
Db       2919 VSTPVTQGGTVKVLTQGINTPPVLVHNQLVLTPSIVTTNKKLADPVTLKIETKVLQPANL 2978

Qy       2000 -KNLAPHHASPDPPAPPASASDP--HREKTQSKPF--SIQELELRSLGYHGSSYSPEGVE 2054
Db       2979 GSTLTPHH----PPALPSKLPTEVNHVPSGPSIPADRTVSHLAAAKLDAHSPRPSGPGPS 3034

Qy       2055 PVSPVSSPSLTHDKGL----------------------------PKH--LEELDKSHL- 2082
Db       3035 SFPRASHPSSTASTALSTNATVMLAAGIPVPQFISSIHPEQSVIMPPHSITQTVSLSHLS 3094

Qy       2083 EGELRPKQPG----PVKLGGEAAHLPHLRPLPESQPSSSPLLQTAPGVKGHQRVVTLAQH 2138
Db       3095 QGEVRMNTPTLPSITYSIRPEALHSPR-APL---QP--------------QQIEVRA-- 3133

Qy       2139 ISEVITQDYTRHHPQQLSAPLPAPLYSFPGASCPVLDLRRPPSD---LYLPPPDHGAPAR 2195
Db       3134 -------------PQRASTPQPAP------AGVPALASQHPPEEEVHYHLPVARATAPVQ 3174

Qy       2196 GS------------------------PH----------SEGGKRSPEPNKTSVLGGGEDG 2221
Db       3175 SEVLVMQSEYRLHPYTVPRDVRIMVHPHVTAVSEQPRAADGVVKVPPASKAP----QQPG 3230

Qy       2222 IEPVSPPEGMTEPGHSRSAV-------YPLLYRDGEQTEPSRMGSKSPGNTSQ-----PP 2269
Db       3231 KEAAKTPDAKAAPTPTPAPVPVPVPLPAPAPAPHGE----ARILTVTPSNQLQGLPLTPP 3286

Qy       2270 AFFSKLTESNSAMVKSKKQEINKKLNTHNRNE--PEYNISQPGTEIFNMPAITGTGLMTY 2327
Db       3287 ------------VVVTHGVQI-----VHSSGELFQEYRYGD---------------IRTY 3314

Qy       2328 RSQAVQEH----ASTNMGLEAIIRKALMGKYDQWEE-SPPLSANAFNPLNASASLPAAMP 2382
Db       3315 HPPAQLTHTQFPAASSVGLPSRTKTAAQGPPPEGEPLQPPQPVQSTQPAQPAPPCPPSQ- 3373

Qy       2383 ITAADGRSDHTLTSPGGGGKAK---VSGRPSSRKAKSPAPGLASG--DRPP----SVSSV 2433
Db       3374 -----------LGQPGQPPSSKMPQVSQEAKGTQTGVEQPRLPAGPANRPPEPHTQVQRA 3422

Qy       2434 HSE-GDCNRRTPLTNRVWEDRPSSAGSTPFPYNPLIMRLQAGVMASPPPPGLPAGSGPLA 2492
Db       3423 QAETGPTSFPSPVSVSMKPDLPVSLPTQTAPKQPLFVPTTSG-------PSTPPG---LV 3472

Qy       2493 GPHHAWDEEPK----PLLCSQ 2509
Db       3473 LPHTEFQPAPKQDSSPHLTSQ 3493
```

RESULT 10
US-10-263-929-144
; Sequence 144, Application US/10263929
; Publication No. US20040067535A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jaeseob
; APPLICANT: Galant, Ron
; TITLE OF INVENTION: Alzheimer's Disease Linked Genes
; FILE REFERENCE: LSD-07417
; CURRENT APPLICATION NUMBER: US/10/263,929
; CURRENT FILING DATE: 2002-10-03
; NUMBER OF SEQ ID NOS: 213
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 144
; LENGTH: 3551
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-263-929-144

Query Match 4.1%; Score 547; DB 12; Length 3551;
Best Local Similarity 19.0%; Pred. No. 2.6e-16;
Matches 592; Conservative 346; Mismatches 1117; Indels 1056; Gaps 131;

```
Qy         26 YPVQIARTHTDVGLLEYQHHSRDYASHLSPGSIIQPQRR-RPSLLSEFQPGNE------- 77
Db        543 YEMLTERREERRGSYEYSQERTYYENVRTPGTYPEDSRRDYPARGREFYSEWETYQGEYY 602

Qy         78 --------------RSQELHLRPESHSYLPELGKSEMEFIESKRPR----LELLPDPLLR 119
Db        603 DSRYYDEPREYREYRSDPYEQDIREYSYRQRERERERERFESDRDHERRPIERSQSPVHL 662

Qy        120 PSPLLATGQPAGSEDLTKD--RSL-------TGKLEPVSPP-------------SPPHT 156
Db        663 GRPQSPGVFPAHSERLPSDFERRLYRRSSERSGSCSSVSPPRYDKLEKARLERYTKNEKA 722

Qy        157 DPELELVPPRLSKEELIQNMDRVDREITMVEQQISKLK-----KKQQQLEEEAAKPPEPE 211
Db        723 DKERTFDPERVERERRIVRKEKGEKD--KAERQKRKGKAHSPSSQPSETEQENDREQSPE 780

Qy        212 KPVSPPPIESKHRSLVQIIYDENRKKAEAAH-----------RILEGLGPQVELP----- 255
Db        781 KP----------RGSTKLSRDRADKEGPAKNRLELVPCVVLTRVKEKEGKVIEHPPPEKL 830

Qy        256 -------------LYNQP---------SDTRQYHENIKINQAMRKKLILYFKRRNHA--- 290
Db        831 KARLGRDTTKALALDQKPQAAQGEPAKSDPAR-------GKALREKVL-----PSHAEVG 878

Qy        291 RKQWKQKFCQ--RYDQLMEALEKKVERIENNPRRRAKESKVREYYEKQFPEIRK------ 342
Db        879 EKEGRTKLRKHLKAEQTPELSALDLEKLEARKRRFA-DSGLK--IEKQKPEIKKTSPETE 935

Qy        343 ----------------------------QREL--QERMQSRVGQRGSGLS--------- 362
Db        936 DTRILLKKQPDTSRDGVLLREGESERKPVRKEILKRESKKTKLERLNSALSPKDCQDPAA 995

Qy        363 MSAARSEHEVSEIIDGLSE--------QENLEK-------QMRQLAVI------------ 395
Db        996 VSAGSGSRPSSDVHAGLGELTHGSVETQETQPKKAIPSKPQPKQLQLLENQGPEKEEVRK 1055

Qy        396 -------PPMLYDADQQRIKFINMN---GLMADPMKVYK-----DRQVMNMWSEQEKETF 440
Db       1056 NYCRPREEPAEHRAGQEKPHGGNAEEKLGIDIDHTQSYRKQMEQSRRKQRMEMEIAK--- 1112

Qy        441 REKFMQHPKNFGLIASFLERKTVAEC---------------------VLYYYLTKKNENY 479
Db       1113 AEKFGSPKKD---VDDYERRSLVHEVGKPPQDVTDDSPPSKKRRTDHVDFDICTKRERNY 1169

Qy        480 KS---LVRRSYR-------RRGK---------SQQQQQQQQQQQQQQQQQPMPRSSQEEK 520
Db       1170 RSSRQISEDSERTSCSPSVRHGSFHDDDDPRGSPRLVSVKGSPKGDEKGLPYPNAAVRDD 1229

Qy        521 -------DEKEKEKEAEKEEEK--------------PEVEND-------------KEDLL 546
Db       1230 PLKCNPYDSGKREQTADTAKIKLSVLNSEGEPSRWDPPMKQDPSRFDVSFPNSVIKRDSL 1289

Qy        547 KEKT--DDTSGE---DNDEKEAVASKGRKTAN---------------SQGRRKGRITRS 585
Db       1290 RKRSVRDLEPGEVPSDSDEDAEHRSQSPRASSFYDSPRLSFLLRDRDQKLRERDERLASS 1349

Qy        586 MANEANS------EEAITPQQSAELASMELNESSRWTEEEMETAKKGLLEHGRNWSAIAR 639
Db       1350 L--ERNKFYSFALDKTITPDTKALLERAKSLSSSR---EE-------------NWSFL-- 1389

Qy        640 MVGSKTVSQCKNFYFNYKKRQNLDE--------ILQQHKLKMEKE------RNARR---- 681
Db       1390 ----DWDSRFANFR-NNKDKEKVDSAPRPIPSWYMKKKKIRTDSEGKLADKKDERREEEQ 1444

Qy        682 --------------------KKKKAPAAASEEAAFPPVVEDEEMEASGVSGNEEEMVEEA 721
Db       1445 ERQELFASRFLHSSIFEQDSKRLQYLERKSEESDLPPGGLYGRQASEGANSTSDSVQEPV 1504

Qy        722 EALHASGNEVPRGECSGPATVNNSSDTESIPSPHT-----EAAKDTGQNGPKPPATLGAD 776
Db       1505 VLFHSRFMELTRMRQKEKEKDQKPKEAEKQEEPETHPKTPEPAAETKEPEPKAPVSAGLP 1564

Qy        777 GPPPGPPTP------PRRTSRAPIEPTPASEATGAPTP-------------PPAPPSPS 816
```





RESULT 11
US-09-864-761-34248
; Sequence 34248, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

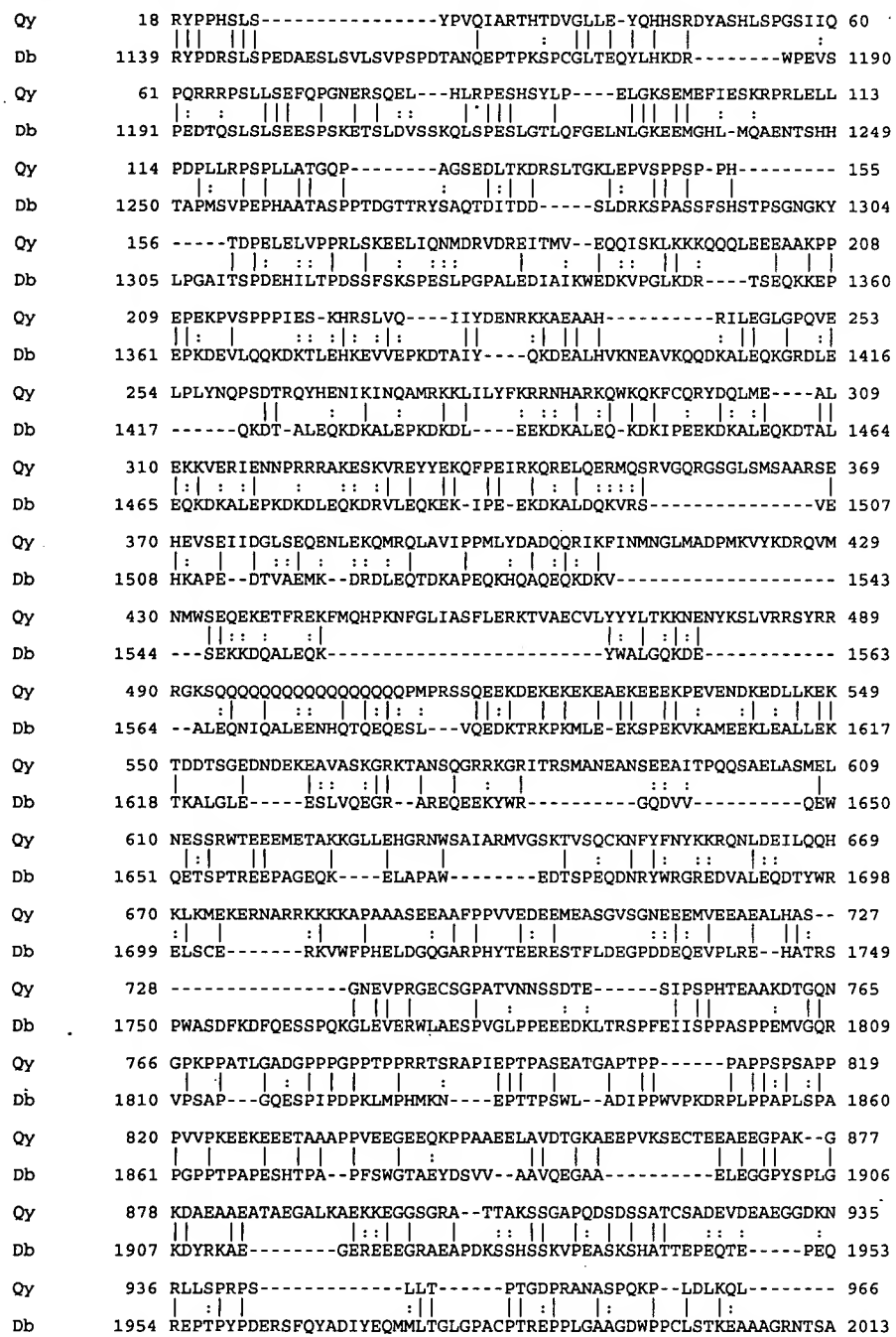
```
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34248
;   LENGTH: 2665
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL034555.2
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL =  10
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 8.9
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.8
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL =  14
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 7.2
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 9.5
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 7.1
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 9.3
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 7.7
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL =  12
;   OTHER INFORMATION: EST_HUMAN HIT: AU117052.1, EVALUE 0.00e+00
;   OTHER INFORMATION: SWISSPROT HIT: P08640, EVALUE 3.00e-10
US-09-864-761-34248

  Query Match             3.9%;  Score 518;  DB 9;  Length 2665;
  Best Local Similarity   19.3%;  Pred. No. 4e-15;
  Matches  505;  Conservative  315;  Mismatches  863;  Indels  934;  Gaps  120;

Qy         71 EFQPGNERSQELHLRPE-------SHSYLPELGKS------EMEFIESKR---PRLELLP 114
              | : | |:|   ::    :       : ||  :: :|      |||  :|::   |: ::
Db        525 ERKSGQEKSHSVNTEEKIGIDIDHTQSYRKQMEQSRRKQQMEMEIAKSEKFGSPKKDV-- 582

Qy        115 DPLLRPSPLLATGQPAGSEDLTKDRSLTGKLEPVSPPSPPHTDPELELVPPRLSKEELIQ 174
              |   | | :   |:|   :|:| |          ||||                   :
Db        583 DEYERRSLVHEVGKP--PQDVTDD----------SPPSKK------------------K 611

Qy        175 NMDRVDREI-TMVEQQISKLKKKQQQLEEEAAKPP-------EPEKPVSPPPIES----- 221
              || || :| | |:      ::  :  |       |       | | |:  | : |
Db        612 RMDHVDFDICTKRERNYRSSRQISEDSERTGGSPSVRHGSFHEDEDPIGSPRLLSVKGSP 671

Qy        222 --------------KHRSLVQIIYDENRKK--AEAAHRILEGLGPQVELPLYNQPSDTRQ 265
                          :  ||    || :|::  |: |   |  |  : ||  :    |::
Db        672 KVDEKVLPYSNITVREESLKFNPYDSSRREQMADMAKIKLSVLNSEDELNRW----DSQM 727

Qy        266 YHENIKINQAMRKKLILYFKRRNHARKQWKQKFCQRYDQLMEALEKKVERIENNPRR--- 322
               :  : : :    :|    :|:  ||:  :       ::    ::  |   ::||
Db        728 KQDAGRFDVSFPNSII----KRDSLRKRSVRDL--EPGEVPSDSDEDGEHKSHSPRASAL 781

Qy        323 ----------RAKESKVREYYEKQFPEIRKQRELQERMQSRVGQRGSGL-----SMSAAR 367
                       | :| |:||  |:    : : :     :     :      |     |:|::|
Db        782 YESSRLSFLLRDREDKLRERDERLSSSLERNKFYSFALDKTITPDTKALLERAKSLSSSR 841

Qy        368 SEHEVSEIIDGLSE----QENLEKQMRQLAVIPPMLYDADQQRIKFINMNGLMADPMKVY 423
              |:     :|  |      : | :|:    |  |   :   :::|:  :  | | |  : :
Db        842 EEN--WSFLDWDSRFANFRNNKDKEKVDSAPRPIPSWYMKKKKIR-TDSEGKMDDKKEDH 898

Qy        424 KDRQVMNMWSEQEKETFREKFMQHPKNFGLIASFLERKTVAECVLYYYLTKKNENYKSLV 483
              |: :       :: :| |  :|: |   |    :  |:                :| :| |:   :
Db        899 KEEE------QERQELFASRFL-HSSIFEQDSKRLQ----------HLERKEEDSDFIS 940

Qy        484 RRSYRRRGKSQQQQQQQQQQQQQQQQP----------MPRSSQEEKDEKEKEKEAEKEE 533
              | |   ||   |  :          |:|          : |  |:|:: :| || ||:|
Db        941 GRIY---GK---QTSEGANSTTDSIQEPVVLFHSRFMELTRMQQKEKEKDQKPKEVEKQE 994

Qy        534 EKPEVENDKEDLLKEKTDDTSGEDNDEKEAVASKGRKTANSQGRRKGRITRSMANEANSE 593
               : ||         || ::: |: |                                :|  :
Db        995 ---DTEN------HPKTPESAPENKD--------------------------SELKTP 1017

Qy        594 EAITPQQSAELASMELNESSRWTEEEMETAKKGLLEHGRNWSAIARMVGSKTVSQCKNFY 653
              :: |                  |   :|:|           ||: :  | |||
Db       1018 PSVGPPSV-------------TVVTLESAP----------SALEKTTGDKTV------- 1046

Qy        654 FNYKKRQNLDEILQQHKLKMEKERNARRKKKKAPAAASEEAAFPPVVEDEEMEASGVSGN 713
                                          :||    |:   |  | :|   ||  :
Db       1047 ------------------------------EAPLVTEEKTVEPATVSEEAKPASEPAPA 1075

Qy        714 EEEMVEEAEALHASGNEVPRGECSGPATVNNSSDTESIPSPHTEAAKDTGQNGPKPPATL 773
               | :|:          ::| |             | |:   |   |  : | :| :||  |
Db       1076 PVEQLEQV--------DLPPGA---------DPDKEAAMMP---AGVEEGSSGDQPP-YL 1114

Qy        774 GADGPPPGPPTPPRRTSRAP--IEPTPASEATGAPTPPPAPPSPSAPPPVVPKEEKEEET 831
               |     ||||     |:|   ::| | |       | | : |:  :      || || : |
Db       1115 DAK-----PPTPGASFSQAESNVDPEPDS------TQPLSKPAQKSEEANEPKAEKPDAT 1163

Qy        832 AAAPP-----VEEGEEQKPPAAEELAVD---TGKAEEPVKSECTEEAEEGPAKGKDAEAA 883
               | | |        |   | :|||:|:| ||     | ::| ||: ::   :  |      :
Db       1164 ADAEPDANQKAEAAPESQPPASEDLEVDPPVAAKDKKPNKSKRSKTPVQAAAVSIVEKPV 1223

Qy        884 EATAEGALKAEKKEGGSGRATTAKSSGAPQDSDS-SATCSADEVDEAEGGDKNRLLSPRP 942
               :|   : : |   | |     |       :::  : | | :   : |          | |
Db       1224 TRKSERIDREKLKRSNSPRGEAQKLLELKMEAEKITRTASKNSAADLE--------HPEP 1275

Qy        943 SL-LTPT------------GDPRANASPQK-PLD--------LKQLKQRAAAIPPIQVTK 980
              || |: |              ||   | || | |::       |::  | |||:|    :
Db       1276 SLPLSRTRRRNVRSVYATMGD-HENRSPVKEPVEQPRVTRKRLERELQEAAAVP--TTPR 1332

Qy        981 VHEPPR--------EDAAPTKPAPPAPPP----------------PQNL----QPESDAP 1012
               ||:         |:    :||    ||                ||     :|: ||
Db       1333 RGRPPKTRRRADEEEENEAKEPAETLKPPEGWRSPRSQKTAAGGGPQGKKGKNEPKVDAT 1392

Qy       1013 QQPGSSPRG-----KSRSPAPPADKEAFAAEAQKLPGDPPCWTSGLPFPVPPREVIKASP 1067
              :   ::  |     |  | | | | :|   :| ::   |     |     |    | ||::  |
Db       1393 RPEATTEVGPQIGVKESSMEPKAAEEEAGSEQKRDRKD--AGTDKNPPETAPVEVVEKKP 1450

Qy       1068 HAPDPSAFSYAPPGHPLPLGLHDTARPVLPRPPTISNPPPLISSAKHPSVLERQIGAISQ 1127
               ||: :: |          |   :|  : :  :: |    : :|  |      | |
Db       1451 -APEKNSKS--------KRGRSRNSRLAVDKSASLKN----VDAAVSPRGAAAQAGERES 1497

Qy       1128 GMSVQLHVPYSEHAKAPVGPVTMGLPLPMDP-----KKLAPFSG--------VKQEQL-- 1172
```

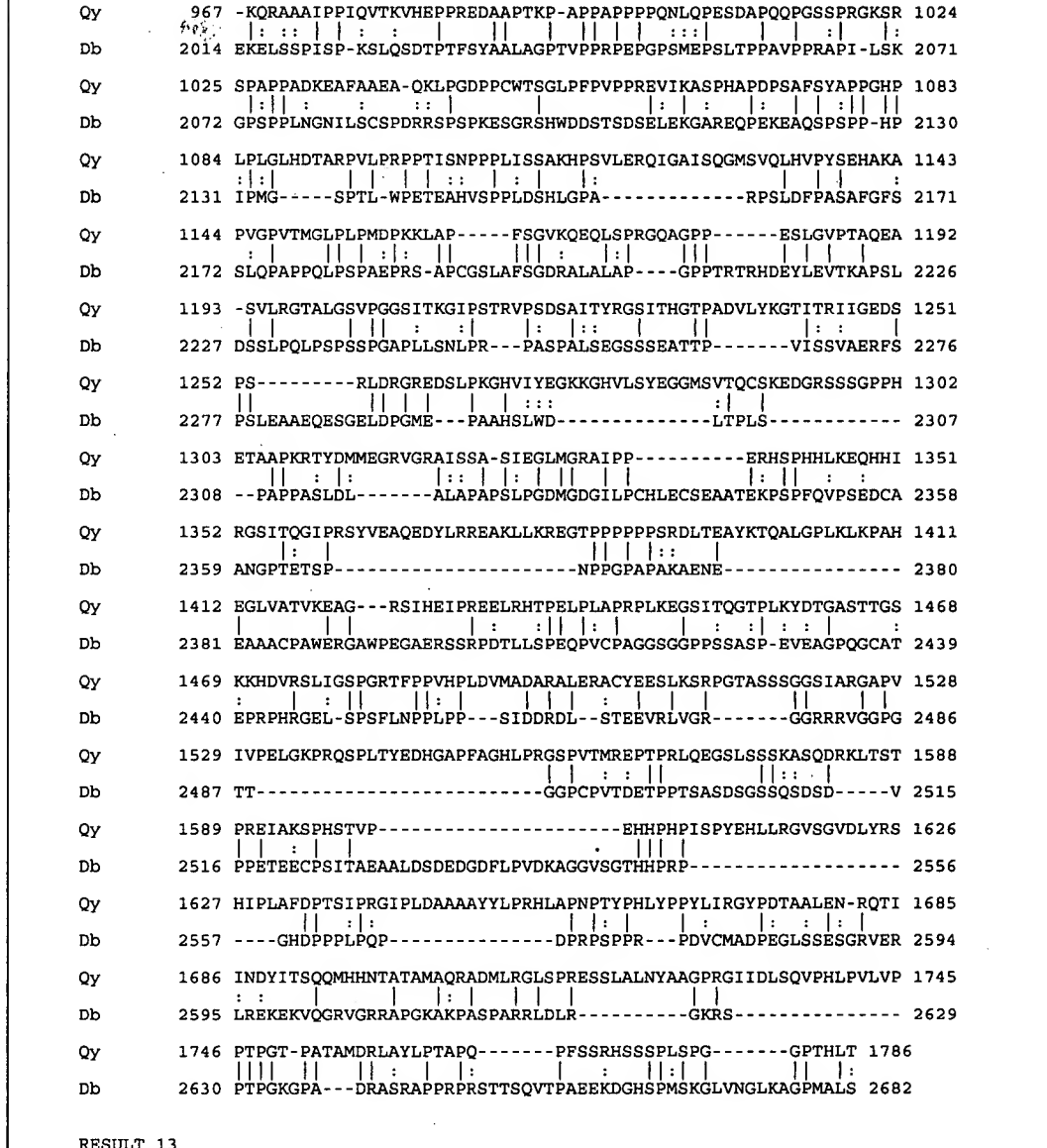
```
Db       1498 GV-VAVSPEKSESPQKEDGLSSQLKSDPVDPDKEPEKEDVSASGPSPEATQLAKQMELEQ 1556

Qy       1173 ----------SPRGQAGPPESLGVPTAQEASVLRGTALGSV--------- 1203

Db       1557 AVEHIAKLAEASASAAYKADAPEGLA--PEDRDKPAHQASETELAAAIGSIINDISGEPE 1614

Qy       1204 --------PGGSITKGIPSTRV----PSDSAITYRGSITHGTPADVLYKGTI-TRIIGED 1250

Db       1615 NFPAPPPYPGESQTDLQPPAGAQALQPSE----------EGMETDEAVSGILETEAATES 1664

Qy       1251 S------------PSRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSS 1298

Db       1665 SRPPVNAPDPSAGPTDTKEARGNSSETSHSVPEAKGSK------EVEVTLVRKDKGRQ-- 1716

Qy       1299 GPPHETAAPKRTYDMMEGRVGRAISSASIEGLMGRAIPPERHSPHHLKEQHHIRG---SI 1355

Db       1717 ---------KTTRSRRKRNTNKKV-----------VAPVESHVP----ESNQAQGESPAA 1752

Qy       1356 TQGIPRSYVEAQEDYLRREAKLLKREGTPPPPPPSRDLTEAYKTQALGPLKLKPAHEGLV 1415

Db       1753 NEGTTVQHPEAPQ-----EEKQSEKPHSTPPQSCTSDLSKIPSTE--------------- 1792

Qy       1416 ATVKEAGRSIHEIPREELRHT-----PELPLAPRPLKEGSITQGTPLKYDTGASTTGSKK 1470

Db       1793 -------NSSQEISVEERTPTKASVPPDLPPPPQP---------APVDEEPQA------R 1830

Qy       1471 HDVRSLIGSPGRTFP--PVHPLDVMADARALERACYEESLKSRPGTASSSGGSIARGAPV 1528

Db       1831 FRVHSIIESDPVTPPSDPSIPIPTLPSV----------------TAAKLSPPVASG--- 1870

Qy       1529 IVPELGKPRQSP-------LTYEDHGAPFAGHLPRGSPVTMREPTPRLQEGSLSS---SK 1578

Db       1871 -----GIPHQSPPTKVTEWITRQE------------EPRAQSTPSPALPPDTKASDVDTS 1913

Qy       1579 ASQDRKLTSTPREIAKS-------------PHSTVPEHHPHPISPYEHLLRGVSGVDLYR 1625

Db       1914 SSTLRKILMDPKYVSATSVTSTSVTTAIAEPVSAAPCLHEAPPPPVD------SKKPLEE 1967

Qy       1626 SHIPLAFDPTSIPRGIPLDAAAAYYLPRHLAPNPTYPHLYPPYLIRGYPDTAALENRQTI 1685

Db       1968 KTAPPVTNNSEIQASEVLVAADKEKVAPVIAPKIT-------SVISRMPVSIDLENSQKI 2020

Qy       1686 INDYITSQQMHHNTATAMAQRADMLRGLSPRESSLALNYAAGPRGIIDLSQVPHLPVLVP 1745

Db       2021 --------------TLAKPAPQTLTGL-----------VSALTGLVNVSLVP-VNALKG 2053

Qy       1746 PTPGTPATAMDRLAYLPTAPQPFSSRHSSSPLSPGGPTHLTKPTTTSSSERERDRDRERD 1805

Db       2054 PVKGSVTTLKSLVS-----------------TPAGPVNVLKGPV--------------- 2080

Qy       1806 RDREREKSILTSTTTVEHAPI--------WRPGTEQSSGS----------------SGSS 1841

Db       2081 -------NVLTGPVNVLTTPVNATVGTVNAAPGTVNAAASAVNATASAVTVTAGAVTAAS 2133

Qy       1842 GGGGGS----------SSRPASHSHAHQHSPISPRTQDALQQRPSVLHNTGMKGIIT 1888

Db       2134 GGVTATTGTVTMAGAVIAPSTKCKQRASANENSRFHPGSMPVIDDRPA---DAG-SGAGL 2189

Qy       1889 AVEPSKPTVLRSTS---TSSPVRPAA---TFPPATHCPLGGTLDGVYPTLMEPVLLPKEA 1942

Db       2190 RVNTSEGVVLLSYSGQKTEGPQRISAKISQIPPAS------AMDIEFQQSVSKSQVKPDS 2243

Qy       1943 PRVARP--ERPRADTGHAFLA------------KPPARSGLEPASSPSKGSEPRPL--- 1984

Db       2244 VTASQPPSKGPQAPAGYANVATHSTLVLTAQTYNASPVISSVK-ADRPSL-EKPEPIHLS 2301

Qy       1985 ------------------VPPVSGHATIARTPA------------------------ 1999

Db       2302 VSTPVTQGGTVKVLTQGINTPPVLVHNQLVLTPSIVTTNKKLADPVTLKIETKVLQPANL 2361

Qy       2000 -KNLAPHHASPDPPAPPASASDP--HREKTQSKPF--SIQELELRSLGYHGSSYSPEGVE 2054

Db       2362 GSTLTPHH----PPALPSKLPTEVNHVPSGPSIPADRTVSHLAAAKLDAHSPRPSGPGPS 2417

Qy       2055 PVSPVSSPSLTHDKGL----------------------------PKH--LEELDKSHL- 2082

Db       2418 SFPRASHPSSTASTALSTNATVMLAAGIPVPQFISSIHPEQSVIMPPHSITQTVSLSHLS 2477

Qy       2083 EGELRPKQPG----PVKLGGEAAHLPHLRPLPESQPSSSPLLQTAPGVKGHQRVVTLAQH 2138

Db       2478 QGEVRMNTPTLPSITYSIRPEALHSPR-APL---QP--------------QQIEVRA-- 2516

Qy       2139 ISEVITQDYTRHHPQQLSAPLPAPLYSFPGASCPVLDLRRPPSD---LYLPPPDHGAPAR 2195

Db       2517 ------------PQRASTPQPAP------AGVPALASQHPPEEEVHYHLPVARATAPVQ 2557

Qy       2196 GS----------------------PHSEGGKRSPEPNKTSVLGGGEDGIEPVSPPEGM 2231

Db       2558 SEVLVMQSEYRLHPYTVPRDVRIMVHPHVTAVSEQPR---------AADGVVKVPPASKA 2608

Qy       2232 TEPGHSRSAVYPLLYRDGEQTEPSRMGSKSPGNTSQP 2268

Db       2609 PQ------------------QPGKEAAKTPDAKAAP 2626
```

RESULT 12
US-10-415-187-5
; Sequence 5, Application US/10415187
; Publication No. US20040044184A1
; GENERAL INFORMATION:
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: YAO, Monique G.
; APPLICANT: WALIA, Narinder K.
; APPLICANT: GIETZEN, Kimberly J.
; APPLICANT: THANGAVELU, Kavitha
; APPLICANT: LU, Yan
; APPLICANT: DING, Li
; APPLICANT: YUE, Henry
; APPLICANT: TANG, Y. Tom
; APPLICANT: LAL, Preeti G.
; APPLICANT: BATRA, Sajeev
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: SANJANWALA, Madhu S.
; APPLICANT: ARVIZU, Chandra
; APPLICANT: RAMKUMAR, Jayalaxmi
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: GURURAJAN, Rajagopal
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: XU, Yuming
; APPLICANT: BURFORD, Neil
; TITLE OF INVENTION: CYTOSKELETON-ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0828 USN
; CURRENT APPLICATION NUMBER: US/10/415,187
; CURRENT FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: PCT/US01/50983
; PRIOR FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: US 60/244,022
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US 60/247,370
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: US 60/251,831
; PRIOR FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PERL Program
; SEQ ID NO 5
; LENGTH: 2803
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20040044184A1 5844189CD1
US-10-415-187-5

Query Match 3.7%; Score 485.5; DB 12; Length 2803;
Best Local Similarity 21.3%; Pred. No. 1.3e-13;

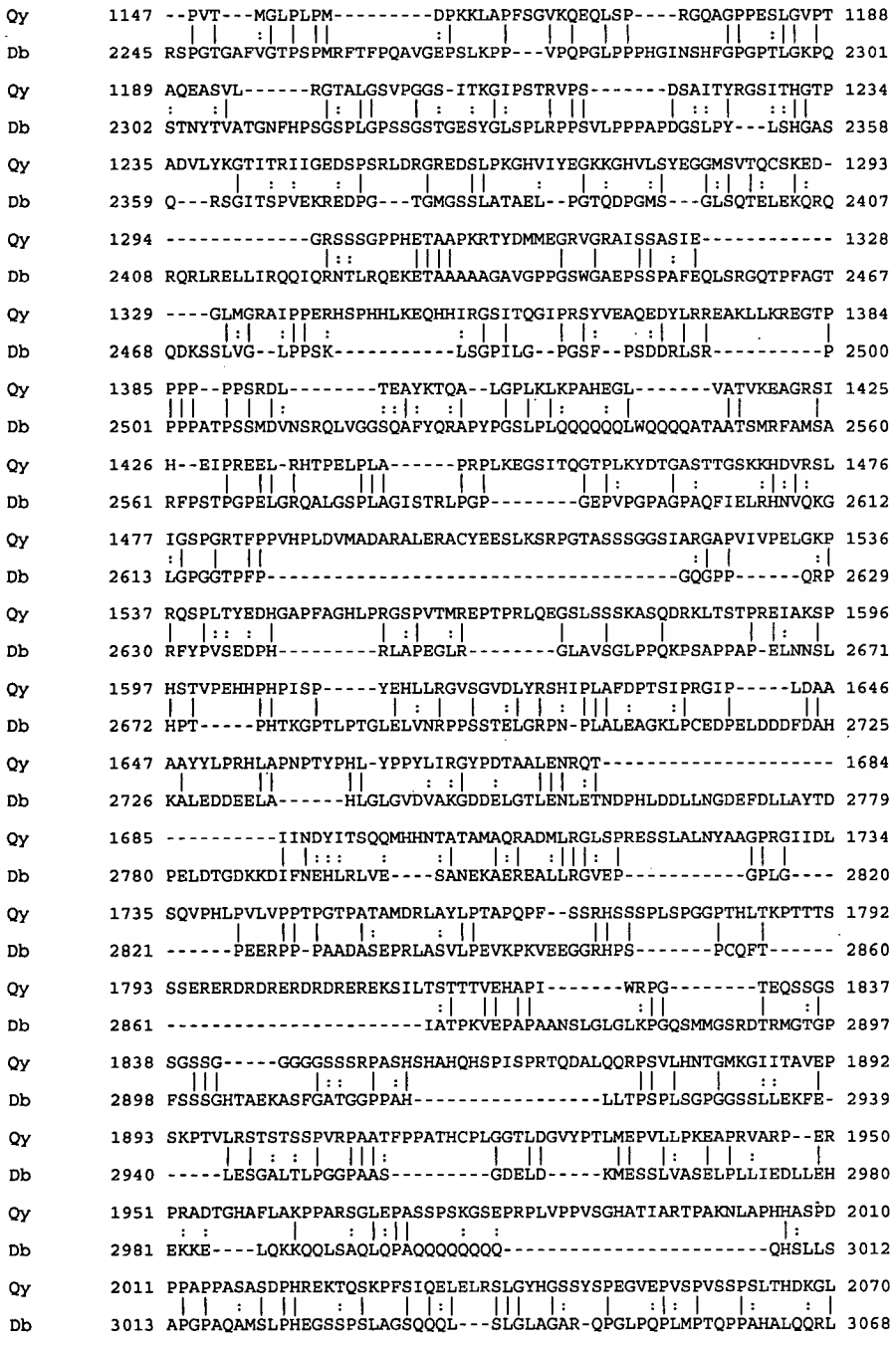
Matches 420; Conservative 211; Mismatches 706; Indels 639; Gaps 99;

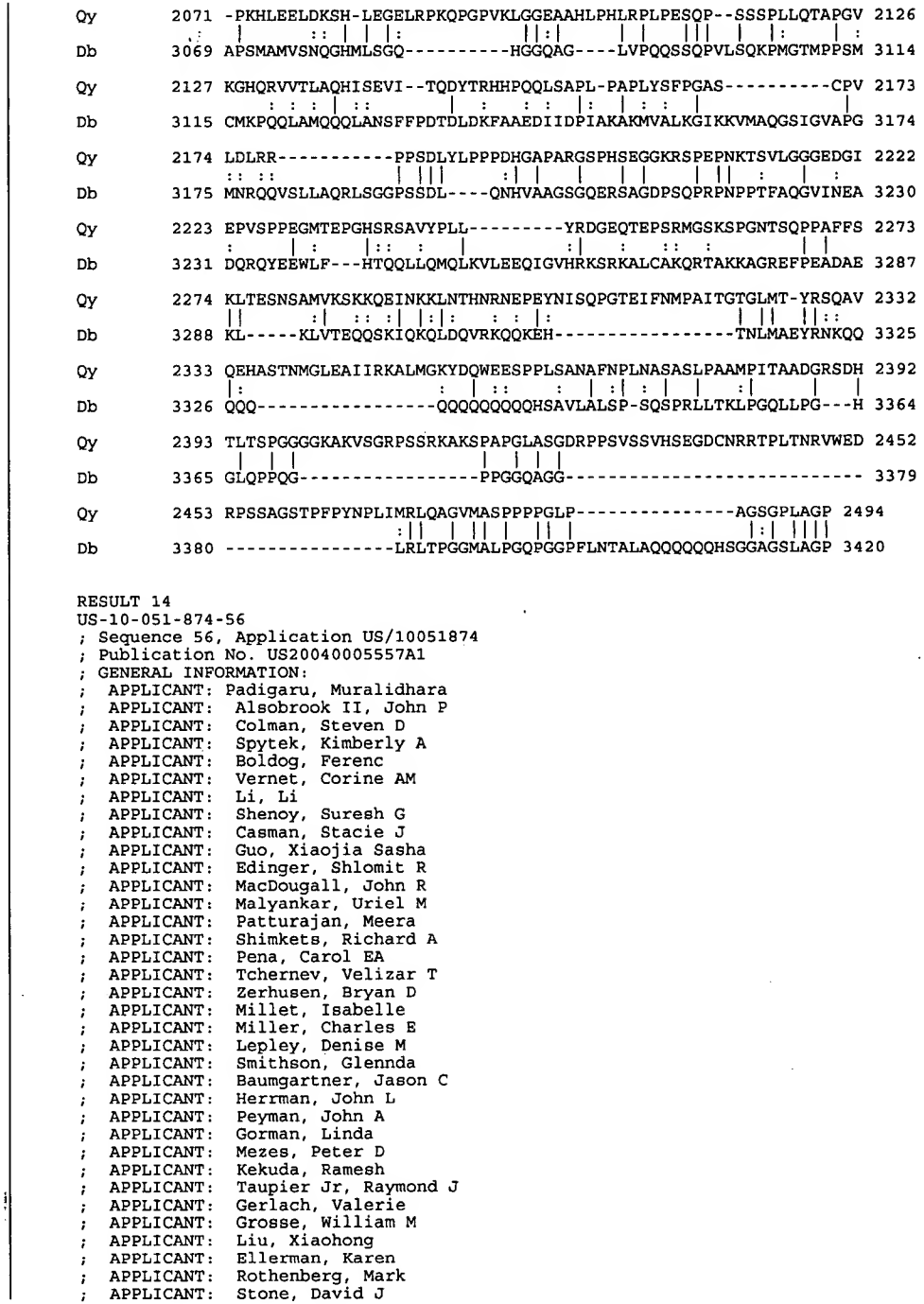




RESULT 13
US-10-051-874-167
; Sequence 167, Application US/10051874
; Publication No. US20040005557A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Alsobrook II, John P
; APPLICANT: Colman, Steven D
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Boldog, Ferenc
; APPLICANT: Vernet, Corine AM

```
;  APPLICANT:  Li, Li
;  APPLICANT:  Shenoy, Suresh G
;  APPLICANT:  Casman, Stacie J
;  APPLICANT:  Guo, Xiaojia Sasha
;  APPLICANT:  Edinger, Shlomit R
;  APPLICANT:  MacDougall, John R
;  APPLICANT:  Malyankar, Uriel M
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Shimkets, Richard A
;  APPLICANT:  Pena, Carol EA
;  APPLICANT:  Tchernev, Velizar T
;  APPLICANT:  Zerhusen, Bryan D
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  Miller, Charles E
;  APPLICANT:  Lepley, Denise M
;  APPLICANT:  Smithson, Glennda
;  APPLICANT:  Baumgartner, Jason C
;  APPLICANT:  Herrman, John L
;  APPLICANT:  Peyman, John A
;  APPLICANT:  Gorman, Linda
;  APPLICANT:  Mezes, Peter D
;  APPLICANT:  Kekuda, Ramesh
;  APPLICANT:  Taupier Jr, Raymond J
;  APPLICANT:  Gerlach, Valerie
;  APPLICANT:  Grosse, William M
;  APPLICANT:  Liu, Xiaohong
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  Rothenberg, Mark
;  APPLICANT:  Stone, David J
;  APPLICANT:  Burgess, Catherine E
;  TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
;  TITLE OF INVENTION:  USING THE SAME
;  FILE REFERENCE: 21402-245
;  CURRENT APPLICATION NUMBER: US/10/051,874
;  CURRENT FILING DATE:  2002-09-25
;  PRIOR APPLICATION NUMBER: 60/268,595
;  PRIOR FILING DATE: 2001-02-14
;  PRIOR APPLICATION NUMBER: 60/325,306
;  PRIOR FILING DATE: 2001-09-27
;  PRIOR APPLICATION NUMBER: 60/262,587
;  PRIOR FILING DATE: 2001-01-18
;  PRIOR APPLICATION NUMBER: 60/272,409
;  PRIOR FILING DATE: 2001-02-28
;  PRIOR APPLICATION NUMBER: 60/262,454
;  PRIOR FILING DATE: 2001-01-18
;  PRIOR APPLICATION NUMBER: 60/276,777
;  PRIOR FILING DATE: 2001-03-16
;  PRIOR APPLICATION NUMBER: 60/291,672
;  PRIOR FILING DATE: 2001-05-17
;  PRIOR APPLICATION NUMBER: 60/330,336
;  PRIOR FILING DATE: 2001-10-18
;  PRIOR APPLICATION NUMBER: 60/265,530
;  PRIOR FILING DATE: 2001-01-31
;  PRIOR APPLICATION NUMBER: 60/261,376
;  PRIOR FILING DATE: 2001-01-16
;  NUMBER OF SEQ ID NOS: 269
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 167
;   LENGTH: 5262
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-051-874-167

  Query Match             3.6%;  Score 472.5;  DB 15;  Length 5262;
  Best Local Similarity   20.1%;  Pred. No. 1.1e-12;
  Matches  555;  Conservative  286;  Mismatches  887;  Indels 1029;  Gaps  142;

Qy        145 LEPVSPPSPPHTDPELELVPPRLSKEE----------LIQNMDRVDREITMVEQQISKLK 194
              ::||:| :||      |||| :: : |          | :    : | :||     | |
Db       1286 VKPVAPVAPP------ELVPMKVKEPEPQYFRFEGVWLTETGMALLRNLTM-----SPLH 1334

Qy        195 KKQQQ-----LEEEAA-KPPEPEKPVSPPPIESKHRSLVQIIYDENRKKAEAAHRILEG- 247
              |::|:     |  ||  :  ||   : |   :       :::  |   :       |||

Db       1335 KRRQRRGRLGLPGEAGLEGSEPSDALGPDDKKDGDLDTDELLKGEGGVEHMECEIKLEGP 1394

Qy        248 LGPQVELPLYNQPSDTRQYHENIKINQAMRKKL--------ILYFKRRNHARKQWKQKFC 299
              : | || |   :  ::::           ||:          :  :|::| | : |
Db       1395 VSPDVE-PGKEETEESKK----------RKRKPYRPGIGGFMVRQRKSHTRTK-KGPAA 1441

Qy        300 QR----------------------YDQLMEALEKKVERIENNPRRRAKESKVREYYEKQ 336
              |                      | |  ||| ::     ||  |:||: : :
Db       1442 QAEVLSGDGQPDEVIPADLPAEGAVEQSLAEGDEKKKQQ-----RRGRKKSKLEDMFPAY 1496

Qy        337 FPEIRKQRELQERMQSR-----VGQRGSGLSMSAARSEHEVSEIIDGLSEQENLEKQMRQ 391
                |  :|| :  :        ||:  ||    |:          || ||:: |     |
Db       1497 LQEAFFGKELLDLSRKALFAVGVGRPSFGLGTPKAKG--------DGGSERKELPTS--Q 1546

Qy        392 LAVIPPMLYDADQQRIK------FINMNGLMADPMKVYKDRQVMNMWSEQEKETFREKFM 445
                | : | : : ::          |: | |:            |: ||
Db       1547 KGDDGPDIADEESRGLEGKADTPGPEDGGVKASPVP----------SDPEK-------- 1587

Qy        446 QHPKNFGLIASFLERKTVAECVLYYYLTKKNENYKSLVRRSYRRRGKSQQQQQQQQQQQ 505
                  |:::| |:| :  |
Db       1588 PGTPGEGMLSSDLDRISTEE---------------------------------------- 1607

Qy        506 QQQQQPMPRSSQEEKDEKEKEKE---AEKEEE---KPEVENDKEDLLK---------- 547
                :|:   | || ::  |:   :|:|:     | :|  :  | :
Db       1608 ------LPK--MESKDLQQLFKDVLGSEREQHLGCGTPGLEGSRTPLQRPFLQGGLPLGN 1659

Qy        548 ----EKTDDTSG------EDNDEKEAVASKGRKTANSQGRRKGRITRSMANEANSEEAIT 597
                    |    |      |: |:    |      :|     |  | |             |
Db       1660 LPSSSPMDSYPGLCQSPFLDSRERGGFFSPEPGEPDSPWTGSGGTTPSTPT--------T 1711

Qy        598 PQQSAELASMELNESS--RWTEEE----METAK---------KGLLEHGRNWSAIARMVG 642
              |    |   :  |: |  || ::|    : |             | :   :||
Db       1712 PTTEGEGDGLSYNQRSLQRWEKDEELGQLSTISPVLYANINFPNLKQDYPDWS------- 1764

Qy        643 SKTVSQCKNFYFNYKKRQNLDEILQQHKLKMEKERNARRKKKKAPAAASEEAAFPPVVED 702
                |:||     ::|    |:     | |    | | |  |    | |:        :
Db       1765 ----SRCKQIMKLWRKVPAADKAPYLQKAK--DNRAAHRINKVQKQAESQ--------IN 1810

Qy        703 EEMEASGVSGNEEEMVEEAEALHASGNEVPRGECSGPATVNNSSDTESIPSPHTEAAKDT 762
              :: :  ::   :        |||      |   | |     :: | | || | |   |
Db       1811 KQTKVGDIARKTDR-----PALHLRIPPQPGALGSPPPA---AAPTIFIGSPTTPAGLST 1862

Qy        763 GQNG-PKPPATLGADGPPPGPPTP-----------------------------PRRTSR 791
              :|  ||||     |  ||| :|                              ||  :
Db       1863 SADGFLKPPA-----GSVPGPDSPGELFLKLPPQVPAQVPSQDPFGLAPAYPLEPRFPTA 1917

Qy        792 APIEPTPASEATGAPTPPPAPPSPSAP----------PPVVPKEEKEEETAAAPPVEEG 840
              | | |    ||||  ||   : | |           ||  |: :          |
Db       1918 PPTYP-PYPSPTGAPAQPPMLGASSRPGAGQPGEFHTTPPGTPRHQPSTPDPFLKP---- 1972

Qy        841 EEQKPPAAEELAVDT------GKAEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAE 894
              : |: : |||         ||| ||: |       |  |:  :          ||: :
Db       1973 ---RCPSLDNLAVPESPGVGGGKASEPLLSP--------PPFGESRK--------ALEVK 2013

Qy        895 KKEGGSGRATTAKSSGAPQ----DSDSSAT------CSADEVDEAEGGDKNRLLSPRPSL 944
              |:| |:    :: | | |     || || |          :| :|        |:|| |
Db       2014 KEELGA----SSPSYGPPNLGFVDSPSSGTHLGGLELKTPDVFKAP-------LTPRASQ 2062

Qy        945 LTPTGDPRANASPQKPLDLKQLKQRAAAIPPIQVTKVHEPPRED----------AAPTK 993
              : |   |     ||:|             || |      |   |           | |
Db       2063 VEPQ-SPGLGLRPQEP-------------PPAQALAPSPPSHPDIFRPGSYTDPYAQPPL 2108

Qy        994 PAPPAPPPPQN---LQPESDAPQQPGS----SPRGKSRSPAP----PADKEAFAAEAQKL 1042
                | ||||::   | | |  | | |    ||: :| | :|     |  |||
Db       2109 TPRPQPPPPESCCALPPRS-LPSDPFSRVPASPQSQSSSQSPLTPRPLSAEAFC------ 2161

Qy       1043 PGDPPCWTSGLPFPVPPREVIKASPHAPDPSAFSYAPPGHPL---PLG-LHDTARPVLPR 1098
                      | || ||       :||| :    ||  |    |   ||  :| |:
Db       2162 -----------PSPVTPRF------QSPDPYS---RPPSRPQSRDPFAPLH---KPPRPQ 2198

Qy       1099 PPTISNPPPLISSAKHPSVLERQIG------AISQGMSVQLHVPYSEHAKAPVG------ 1146
              || :        : |   |:    :|      |:  | : :|      ||| | |
Db       2199 PPEV--------AFKAGSLAHTSLGAGEFPAALPAGPAGEL------HAKVPSGQPPNFV 2244
```





RESULT 14
US-10-051-874-56
; Sequence 56, Application US/10051874
; Publication No. US20040005557A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Alsobrook II, John P
; APPLICANT: Colman, Steven D
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Boldog, Ferenc
; APPLICANT: Vernet, Corine AM
; APPLICANT: Li, Li
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Casman, Stacie J
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Edinger, Shlomit R
; APPLICANT: MacDougall, John R
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard A
; APPLICANT: Pena, Carol EA
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Millet, Isabelle
; APPLICANT: Miller, Charles E
; APPLICANT: Lepley, Denise M
; APPLICANT: Smithson, Glennda
; APPLICANT: Baumgartner, Jason C
; APPLICANT: Herrman, John L
; APPLICANT: Peyman, John A
; APPLICANT: Gorman, Linda
; APPLICANT: Mezes, Peter D
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Gerlach, Valerie
; APPLICANT: Grosse, William M
; APPLICANT: Liu, Xiaohong
; APPLICANT: Ellerman, Karen
; APPLICANT: Rothenberg, Mark
; APPLICANT: Stone, David J

```
;  APPLICANT:  Burgess, Catherine E
;  TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
;  TITLE OF INVENTION:  USING THE SAME
;  FILE REFERENCE: 21402-245
;  CURRENT APPLICATION NUMBER: US/10/051,874
;  CURRENT FILING DATE:  2002-09-25
;  PRIOR APPLICATION NUMBER: 60/268,595
;  PRIOR FILING DATE: 2001-02-14
;  PRIOR APPLICATION NUMBER: 60/325,306
;  PRIOR FILING DATE: 2001-09-27
;  PRIOR APPLICATION NUMBER: 60/262,587
;  PRIOR FILING DATE: 2001-01-18
;  PRIOR APPLICATION NUMBER: 60/272,409
;  PRIOR FILING DATE: 2001-02-28
;  PRIOR APPLICATION NUMBER: 60/262,454
;  PRIOR FILING DATE: 2001-01-18
;  PRIOR APPLICATION NUMBER: 60/276,777
;  PRIOR FILING DATE: 2001-03-16
;  PRIOR APPLICATION NUMBER: 60/291,672
;  PRIOR FILING DATE: 2001-05-17
;  PRIOR APPLICATION NUMBER: 60/330,336
;  PRIOR FILING DATE: 2001-10-18
;  PRIOR APPLICATION NUMBER: 60/265,530
;  PRIOR FILING DATE: 2001-01-31
;  PRIOR APPLICATION NUMBER: 60/261,376
;  PRIOR FILING DATE: 2001-01-16
;  NUMBER OF SEQ ID NOS: 269
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 56
;   LENGTH: 4952
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-051-874-56

  Query Match             3.6%;  Score 470.5;  DB 15;  Length 4952;
  Best Local Similarity   20.3%;  Pred. No. 1.3e-12;
  Matches  558;  Conservative  284;  Mismatches  891;  Indels 1019;  Gaps  144;

Qy        145 LEPVSPPSPPHTDPELELVPPRLSKEE----------LIQNMDRVDREITMVEQQISKLK 194
              ::||:| :||      ||||:: : |          | :   : | :||     | |
Db        981 VKPVAPVAPP------ELVPMKVKEPEPQYFRFEGVWLTETGMALLRNLTM-----SPLH 1029

Qy        195 KKQQQ-----LEEEAA-KPPEPEKPVSPPPIESKHRSLVQIIYDENRKKAEAAHRILEG- 247
              |::|:     |  ||  :  ||   : |   :       :::  |   :       |||
Db       1030 KRRQRRGRLGLPGEAGLEGSEPSDALGPDDKKDGDLDTDELLKGEGGVEHMECEIKLEGP 1089

Qy        248 LGPQVELPLYNQPSDTRQYHENIKINQAMRKKL--------ILYFKRRNHARKQWKQKFC 299
              : | || |   :  ::::           ||:         :  :|::| | : |
Db       1090 VSPDVE-PGKEETEESKK----------RKRKPYRPGIGGFMVRQRKSHTRTK-KGPAA 1136

Qy        300 QR----------------------YDQLMEALEKKVERIENNPRRRAKESKVREYYEKQ 336
              |                      | |  ||| ::     ||  | ||:   :
Db       1137 QAEVLSGDGQPDEVIPADLPAEGAVEQSLAEGDEKKKQQ-----RRGRKRSKLEGMFPAY 1191

Qy        337 FPEIRKQRELQERMQSR-----VGQRGSGLSMSAARSEHEVSEIIDGLSEQENLEKQMRQ 391
                |    :|| :  :       ||:   ||     |:         || ||:: |     |
Db       1192 LQEAFFGKELLDLSRKALFAVGVGRPSFGLGTPKAKG--------DGGSERKELPTS--Q 1241

Qy        392 LAVIPPMLYDADQQRIK------FINMNGLMADPMKVYKDRQVMNMWSEQEKETFREKFM 445
                   | : | : : ::           |: | |:           |: ||
Db       1242 KGDDGPDIADEESRGLEGKADTPGPEDGGVKASPVP----------SDPEK-------- 1282

Qy        446 QHPKNFGLIASFLERKTVAECVLYYYLTKKNENYKSLVRRSYRRRGKSQQQQQQQQQQQ 505
                    |:::| |:| :  |
Db       1283 PGTPGEGMLSSDLDRISTEE------------------------------------- 1302

Qy        506 QQQQQPMPRSSQEEKDEKEKEKE---AEKEEE----KPEVENDKEDLLK---------- 547
                  :|:   | || ::  |:   :|:|:      | :|  :  | :
Db       1303 ------LPK--MESKDLQQLFKDVLGSEREQHLGCGTPGLEGSRTPLQRPFLQGGLPLGN 1354

Qy        548 ----EKTDDTSG------EDNDEKEAVASKGRKTANSQGRRKGRITRSMANEANSEEAIT 597
                    |   |       |: |:   |      :|      | | |            |
Db       1355 LPSSSPMDSYPGLCQSPFLDSRERGGFFSPEPGEPDSPWTGSGGTTPSTPT--------T 1406

Qy        598 PQQSAELASMELNESS--RWTEEE----METAK---------KGLLEHGRNWSAIARMVG 642
              |    |   :  |: |  || ::|    : |             | :   :||
Db       1407 PTTEGEGDGLSYNQRSLQRWEKDEELGQLSTISPVLYANINFPNLKQDYPDWS------- 1459

Qy        643 SKTVSQCKNPYFPNYKKRQNLDEILQQHKLKMEKERNARRKKKKAPAAASEEAAFPPVVED 702
                |:||     ::|    |:     | |     | | |  |     | |:       :
Db       1460 ----SRCKQIMKLWRKVPAADKAPYLQKAK--DNRAAHRINKVQKQAESQ--------IN 1505

Qy        703 EEMEASGVSGNEEEMVEEAEALHASGNEVPRGECSGPATVNNSSDTESIPSPHTEAAKDT 762
              :: :   ::   :       |||      |   | |     :: |  | || | |   |
Db       1506 KQTKVGDIARKTDR-----PALHLRIPPQPGALGSPPPA---AAPTIFIGSPTTPAGLST 1557

Qy        763 GQNG-PKPPATLGADGPPPGPPT--------PPRRTSRA-----------PIEP----- 796
               :|  ||||     |  ||| :        ||:  ::|            |:||
Db       1558 SADGFLKPPA-----GSVPGPDSPGELFLKLPPQVPAQAPSQDPFGLAPAYPLEPRFPTA 1612

Qy        797 ----TPASEATGAPTPPPAPPSPSAP----------PPVVPKEEKEEETAAAPPVEEGE 841
                  |    ||||  ||   : | |           ||  |: :          |
Db       1613 PPTYPPYPSPTGAPAQPPMLGASSRPGAGQPGEFHTTPPGTPRHQPSTPDPFLKP----- 1667

Qy        842 EQKPPAAEELAVDT------GKAEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEK 895
               : |: : |||         ||| ||: |          |  |:  :         ||: :|
Db       1668 --RCPSLDNLAVPESPGVGGGKASEPLLSP--------PPFGESRK--------ALEVKK 1709

Qy        896 KEGGSGRATTAKSSGAPQ----DSDSSAT------CSADEVDEAEGGDKNRLLSPRPSLL 945
              :| |:    :: | | |     || || |          :| :|        |:|| | :
Db       1710 EELGA----SSPSYGPPNLGFVDSPSSGTHLGGLELKTPDVFKAP-------LTPRASQV 1758

Qy        946 TPTGDPRANASPQKPLDLKQLKQRAAAIPPIQVTKVHEPPRED----------AAPTKP 994
              |  |     ||:|           || |     |   |                | |
Db       1759 EPQ-SPGLGLRPQEP----------PPAQALAPSPPSHPDIFRPGSYTDPYAQPPLT 1804

Qy        995 APPAPPPPQN---LQPESDAPQQPGS----SPRGKSRSPAP----PADKEAFAAEAQKLP 1043
              | ||||::   | | |  |  | |    ||: :| | :|    |  |||
Db       1805 PRPQPPPPESCCALPPRS-LPSDPFSRVPVSPQSQSSSQSPLTPRPLSAEAFC------- 1856

Qy       1044 GDPPCWTSGLPFPVPPREVIKASPHAPDPSAFSYAPPGHPL---PLG-LHDTARPVLPRP 1099
                  | || ||        :||| :    ||  |    |   ||   :| |:|
Db       1857 ----------PSPVTPRF------QSPDPYS---RPPSRPQSRDPFAPLH---KPPRPQP 1894

Qy       1100 PTISNPPPLISSAKHPSVLERQIGAISQGMSVQLHV-PYSE-HAKAPVG--------PVT 1149
              | :       : |  |:    :||     |     |  | ||| | |          | |
Db       1895 PEV--------AFKAGSLAHTSLGA--GGFPAALPAGPAGELHAKVPSGQPPNFVRSPGT 1944

Qy       1150 ---MGLPLPM---------DPKKLAPFSGVKQEQLSP----RGQAGPPESLGVPTAQEAS 1193
                :| | ||          :|    |  | | |          ||  :|| | :   :
Db       1945 GAFVGTPSPMRFTFPQAVGEPSLKPP---VPQPGLPPPHGINSHFGPGPTLGKPQSTNYT 2001

Qy       1194 VL------RGTALGSVPGGS-ITKGIPSTRVPS-------DSAITYRGSITHGTPADVLY 1239
              |       |: ||   | :  : |:   | ||       | :: |   ::||
Db       2002 VATGNFHPSGSPLGPSSGSTGESYGLSPLRPPSVLPPPAPDGSLPY---LSHGASQ---R 2055

Qy       1240 KGTITRIIGEDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKED------ 1293
              | : :  : |      ||     :  | :  :|   |:| |:  |:
Db       2056 SGITSPVEKREDPG---TGMGSSLATAEL--PGTQDPGMS---GLSQTELEKQRQRQRLR 2107

Qy       1294 --------GRSSSGPPHETAAPKRTYDMMEGRVGRAISSASIE----------------G 1329
                     |::    ||||          | |   || : |
Db       2108 ELLIRQQIQRNTLRQEKETAAAAAGAVGPPGSWGAEPSSPAFEQLSRGQTPFAGTQDKSS 2167

Qy       1330 LMGRAIPPERHSPHHLKEQHHIRGSITQGIPRSYVEAQEDYLRREAKLLKREGTPPPP-- 1387
              |:|  :|| :          : | |    | |:    :| | |           ||||
Db       2168 LVG--LPPSK----------LSGPILG--PGSF--PSDDRLSR----------PPPPAT 2200

Qy       1388 PPSRDL--------TEAYKTQA--LGPLKLKPAHEGL-------VATVKEAGRSIH--EI 1428
              | | |:        ::|: :|   | | |:   : |          ||      |
Db       2201 PSSMDVNSRQLVGGSQAFYQRAPYPGSLPLQQQQQQLWQQQQATAATSMRFAMSARFPST 2260

Qy       1429 PREEL-RHTPELPLA------PRPLKEGSITQGTPLKYDTGASTTGSKKHDVRSLIGSPG 1481
              |  || |     |||       | |:    | |:     :|:|:  :|  |
Db       2261 PGPELGRQALGSPLAGISTRLPGP--------GEPVPGPAGPAQFIELRHNVQKGLGPGG 2312
```

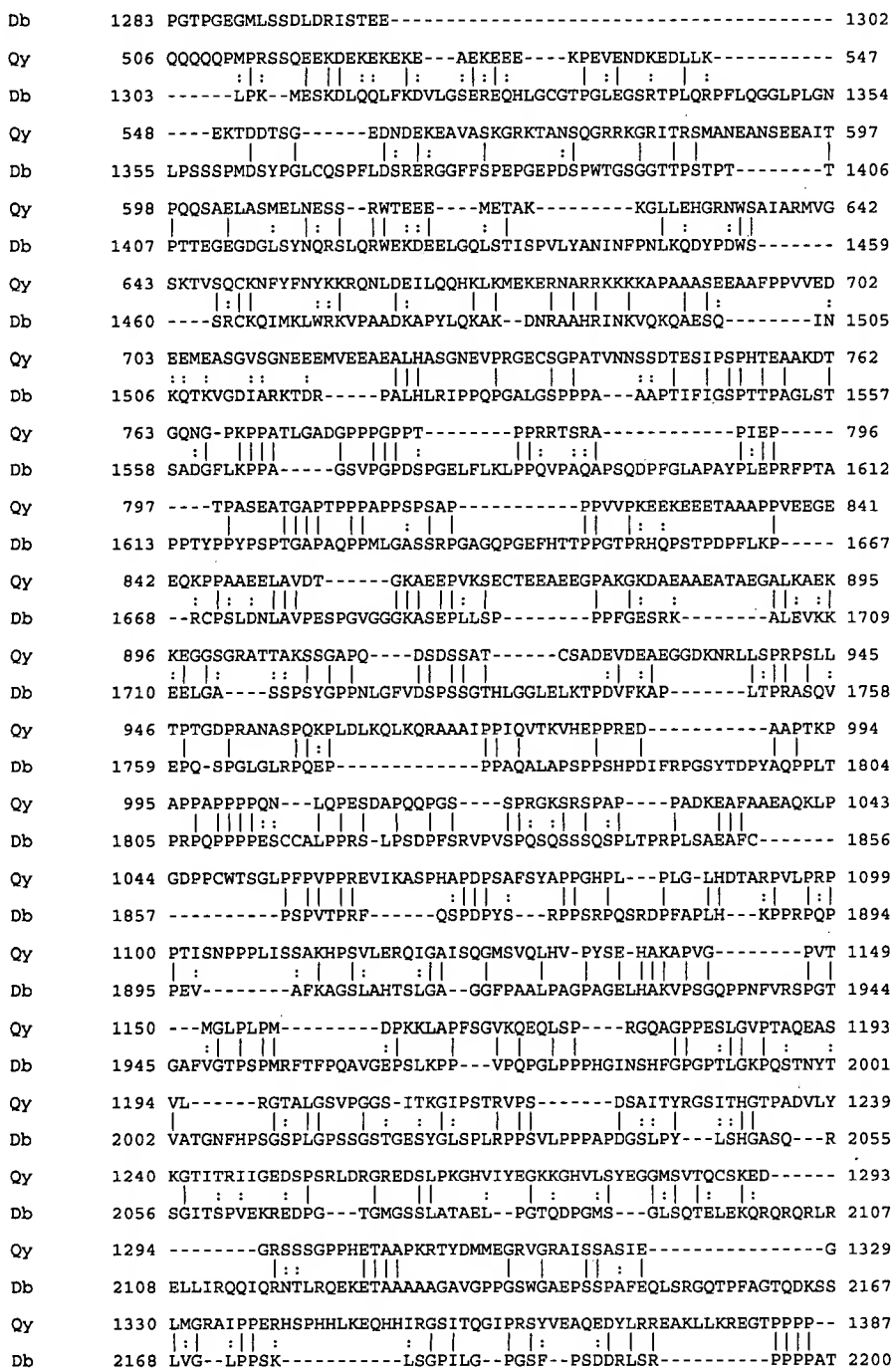
```
Qy       1482 RTFPPVHPLDVMADARALERACYEESLKSRPGTASSSGGSIARGAPVIVPELGKPRQSPL 1541
Db       2313 TPFP-----------------------------------------GQGPP------QRPRFYPV 2329

Qy       1542 TYEDHGAPFAGHLPRGSPVTMREPTPRLQEGSLSSSKASQDRKLTSTPREIAKSPHSTVP 1601
Db       2330 SEDPH---------RLAPEGLR--------GLAVSGLPPQKPSAPPAP-ELNNSLHPT-- 2369

Qy       1602 EHHPHPISP-----YEHLLRGVSGVDLYRSHIPLAFDPTSIPRGIP-----LDAAAAYYL 1651
Db       2370 ---PHTKGPTLPTGLELVNRPPSSTELGRPN-PLALEAGKLPCEDPELDDDFDAHKALED 2425

Qy       1652 PRHLAPNPTYPHL-YPPYLIRGYPDTAALENRQT------------------------- 1684
Db       2426 DEELA------HLGLGVDVAKGDDELGTLENLETNDPHLDDLLNGDEFDLLAYTDPELDT 2479

Qy       1685 -----IINDYITSQQMHHNTATAMAQRADMLRGLSPRESSLALNYAAGPRGIIDLSQVPH 1739
Db       2480 GDKKDIFNEHLRLVE----SANEEAEREALLRGVEP-----------GPLG--------- 2515

Qy       1740 LPVLVPPTPGTPATAMDRLAYLPTAPQPF--SSRHSSSPLSPGGPTHLTKPTTTSSSERE 1797
Db       2516 -PEERPP-PAADASEPRLASVLPEVKPKVEEGGRHPS-------PCQFT----------- 2555

Qy       1798 RDRDRERDRDREREKSILTSTTTVEHAPI-------WRPG--------TEQSSGSSGSSG 1842
Db       2556 -----------------IATPKVEPAPAANSLGLGLKPGQSMMGSRDTRMGTGPFSSSG 2597

Qy       1843 -----GGGGSSSRPASHSHAHQHSPISPRTQDALQQRPSVLHNTGMKGIITAVEPSKPTV 1897
Db       2598 HTAEKASFGATGGPPAH----------------LLTPSPLSGPGGSSLLEKFE------ 2634

Qy       1898 LRSTSTSSPVRPAATFPPATHCPLGGTLDGVYPTLMEPVLLPKEAPRVARP--ERPRADT 1955
Db       2635 LESGALTLPGGPAAS---------GDELD-----KMESSLVASELPLLIEDLLEHEKKE- 2679

Qy       1956 GHAFLAKPPARSGLEPASSPSKGSEPRPLVPPVSGHATIARTPAKNLAPHHASPDPPAPP 2015
Db       2680 ---LQKKQQLSAQLQPAQQQQQQQQQHSLLP----------------------APGPA 2712

Qy       2016 ASASDPHREKTQSKPFSIQELELRSLGYHGSSYSPEGVEPVSPVSSPSLTHDKGL-PKHL 2074
Db       2713 QAMSLPHEGSSPSLAGSQQQL---SLGL-AVARQPGLPQPLMPTQPPAHALQQRLAPSMA 2768

Qy       2075 EELDKSH-LEGELRPKQPGPVKLGGEAAHLPHLRPLPESQP--SSSPLLQTAPGVKGHQR 2131
Db       2769 MVSNQGHMLSGQ----------HGGQAG----LVPQQSSQPVLSQKPMGTMPPSMCMKPQ 2814

Qy       2132 VVTLAQHISEVI--TQDYTRHHPQQLSAPL-PAPLYSFPGAS----------CPVLDLRR 2178
Db       2815 QLAMQQQLANSFFPDTDLDKFAAEDIIGPIAKAKMVALKGIKKVMAQGSIGVAPGMNRQQ 2874

Qy       2179 -----------PPSDLYLPPPDHGAPARGSPHSEGGKRSPEPNKTSVLGGGEDGIEPVSP 2227
Db       2875 VSLLAQRLSGGPSSDL----QNHVAAGSGQERSAGDPSQPRPNPPTFAQGVINEADQRQY 2930

Qy       2228 PEGMTEPGHSRSAVYPLL---------YRDGEQTEPSRMGSKSPGNTSQPPAFFSKLTES 2278
Db       2931 EEWLF---HTQQLLQMQLKVLEEQIGVHRKSRKALCAKQRTAKKAGREFPEADAEKL--- 2984

Qy       2279 NSAMVKSKKQEINKKLNTHNRNEPEYNISQPGTEIFNMPAITGTGLMT-YRSQAVQEHAS 2337
Db       2985 --KLVTEQQSKIQKQLDQVRKQQKEH----------------TNLMAEYRNKQQQQQ-- 3023

Qy       2338 TNMGLEAIIRKALMGKYDQWEESPPLSANAFNPLNASASLPAAMPITAADGRSDHTLTSP 2397
Db       3024 --------------QQQQQQQQQHSAVLALSP-SQSPRLLTKLPGQLLPG---HGLQPP 3064

Qy       2398 GGGGKAKVSGRPSSRKAKSPAPGLASGDRPPSVSSVHSEGDCNRRTPLTNRVWEDRPSSA 2457
Db       3065 QG----------------PPGGQAGG-------------------------------- 3074

Qy       2458 GSTPFPYNPLIMRLQAGVMASPPPPGLP--------------AGSGPLAGP 2494
```

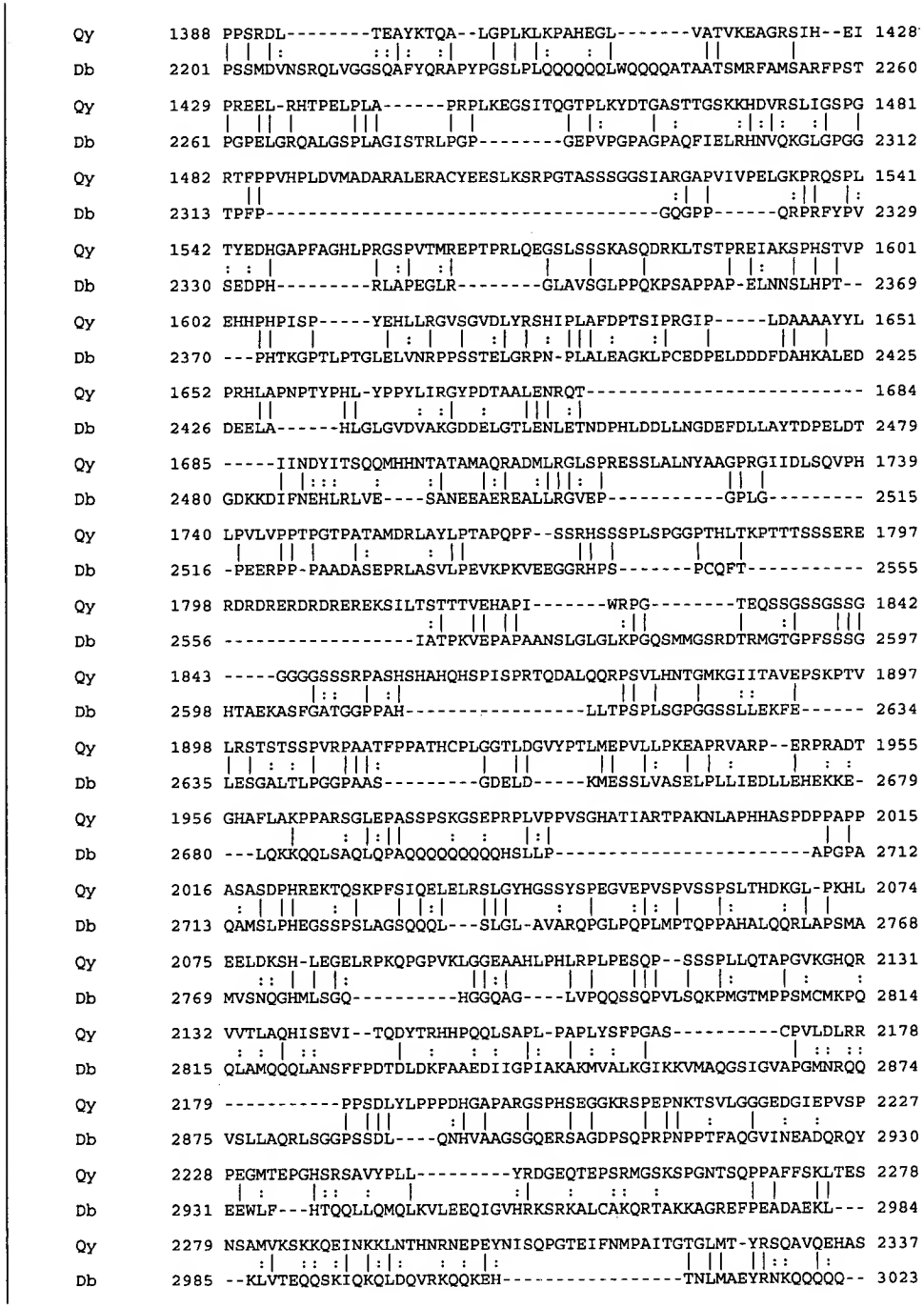
```
Db       3075 ----------LRLTPGGMALPGQPGGPFLNTALAQQQQQQHSGGAGSLAGP 3115
```

RESULT 15
US-10-085-198-112
; Sequence 112, Application US/10085198
; Publication No. US20040009907A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook et al.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-279
; CURRENT APPLICATION NUMBER: US/10/085,198
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/271,646
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 60/276,401
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/311,981
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: 60/312,858
; PRIOR FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: 60/271,840
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/277,324
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/286,096
; PRIOR FILING DATE: 2001-04-21
; PRIOR APPLICATION NUMBER: 60/299,695
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 60/315,614
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/272,405
; PRIOR FILING DATE: 2001-02-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 653
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 112
; LENGTH: 5159
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-085-198-112

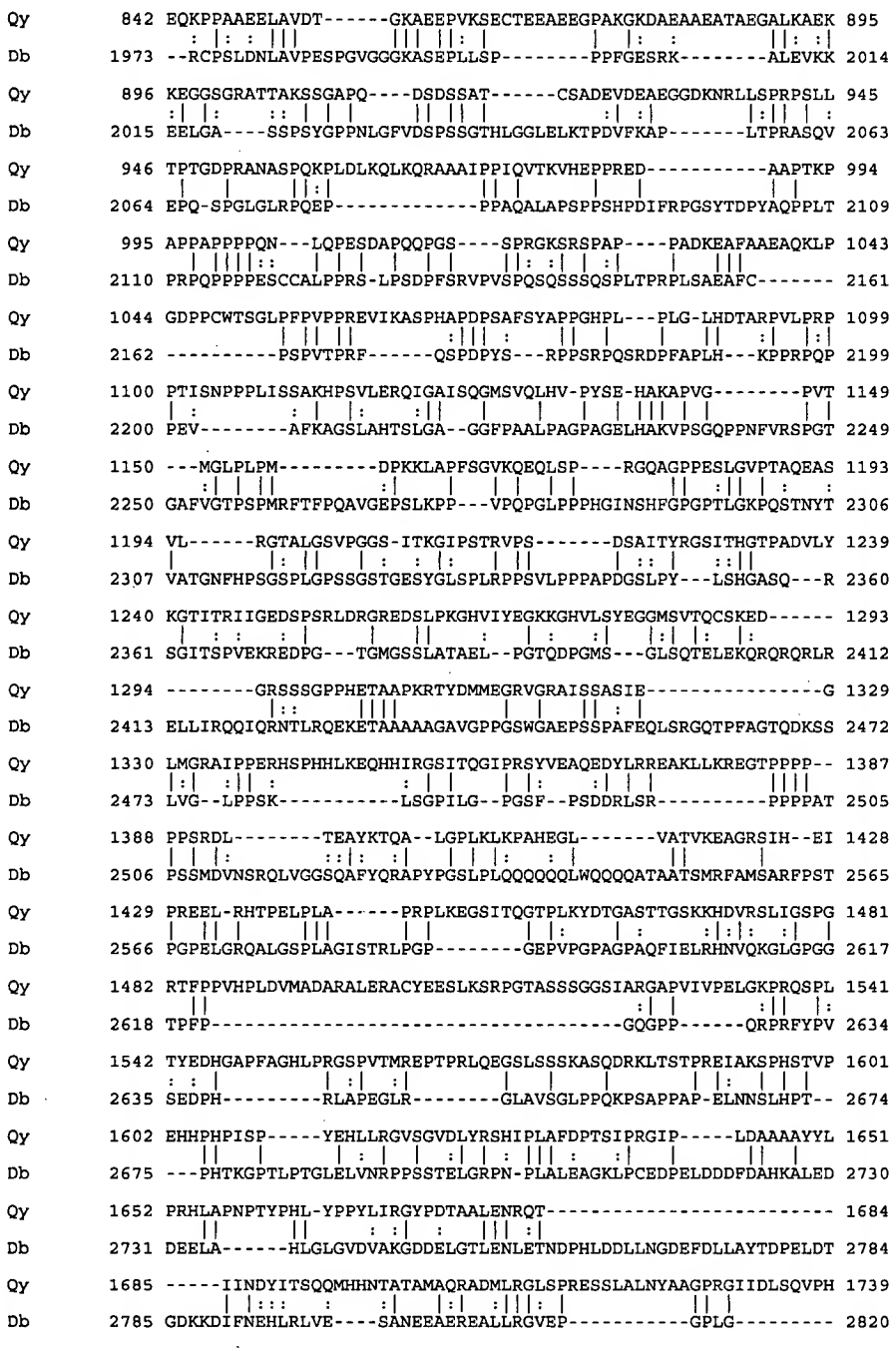
Query Match 3.6%; Score 470.5; DB 15; Length 5159;
Best Local Similarity 20.3%; Pred. No. 1.3e-12;
Matches 558; Conservative 284; Mismatches 891; Indels 1019; Gaps 144;

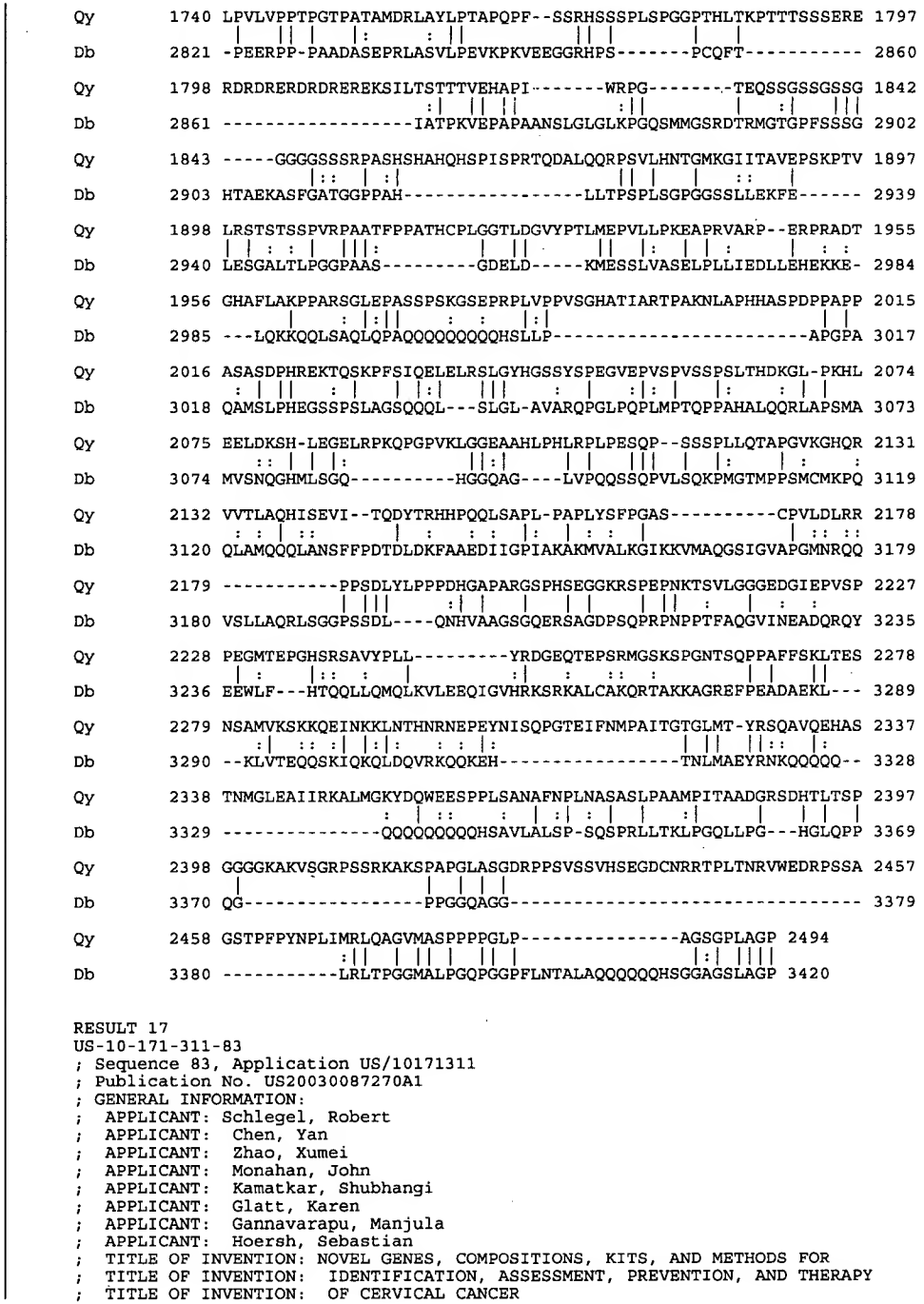
```
Qy        145 LEPVSPPSPPHTDPELELVPPRLSKEE----------LIQNMDRVDREITMVEQQISKLK 194
Db        981 VKPVAPVAPP------ELVPMKVKEPEPQYFRFEGVWLTETGMALLRNLTM-----SPLH 1029

Qy        195 KKQQQ-----LEEEAA-KPPEPEKPVSPPPIESKHRSLVQIIYDENRKKAEAAHRILEG- 247
Db        1030 KRRQRRGRLGLPGEAGLEGSEPSDALGPDDKKDGDLDTDELLKGEGGVEHMECEIKLEGP 1089

Qy        248 LGPQVELPLYNQPSDTRQYHENIKINQAMRKKL--------ILYFKRRNHARKQWKQKFC 299
Db        1090 VSPDVE-PGKEETEESKK----------RKRKPYRPGIGGFMVRQRKSHTRTK-KGPAA 1136

Qy        300 QR----------------------YDQLMEALEKKVERIENNPRRRAKESKVREYYEKQ 336
Db        1137 QAEVLSGDGQPDEVIPADLPAEGAVEQSLAEGDEKKKQQ-----RRGRKRSKLEGMFPAY 1191

Qy        337 FPEIRKQRELQERMQSR-----VGQRGSGLSMSAARSEHEVSEIIDGLSEQENLEKQMRQ 391
Db        1192 LQEAFFGKELLDLSRKALFAVGVGRPSFGLGTPKAKG--------DGGSERKELPTS--Q 1241

Qy        392 LAVIPPMLYDADQQRIK------FINMNGLMADPMKVYKDRQVMNMWSEQEKETFREKFM 445
Db        1242 KGDDGPDIADEESRGLEGKADTPGPEDGGVKASPVP----------SDPEK-------- 1282

Qy        446 QHPKNFGLIASFLERKTVAECVLYYYLTKKNENYKSLVRRSYRRRGKSQQQQQQQQQQQ 505
```





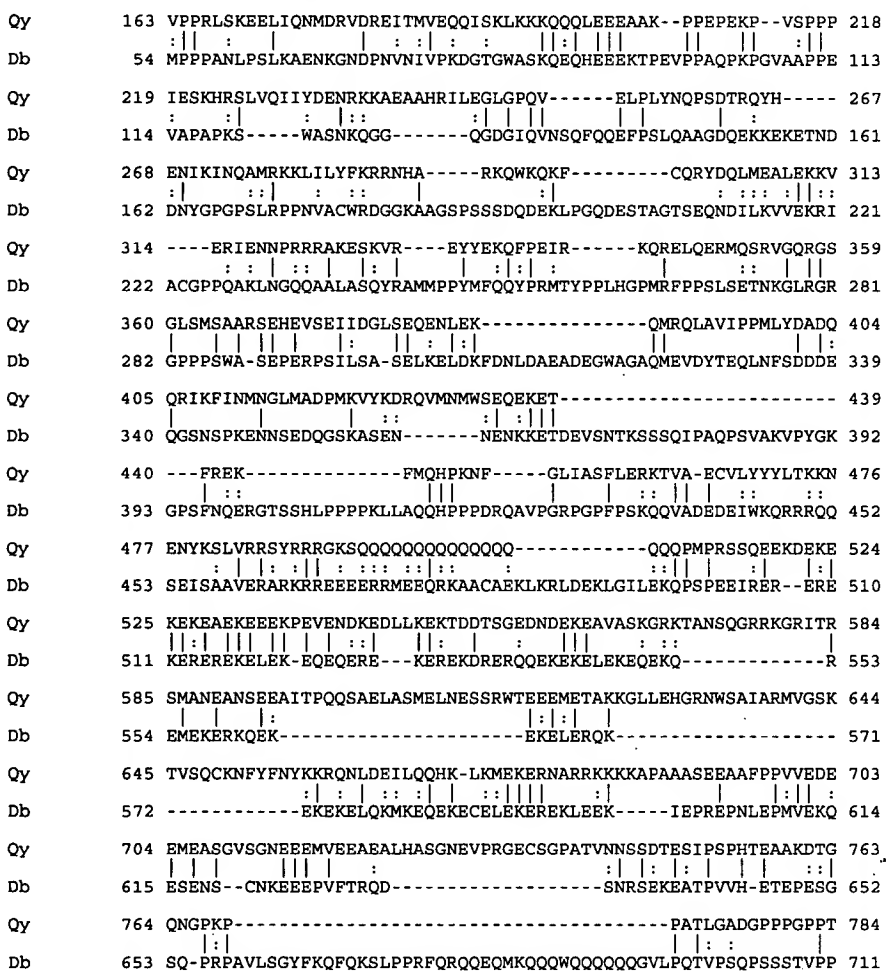
```
Qy       2338 TNMGLEAIIRKALMGKYDQWEESPPLSANAFNPLNASASLPAAMPITAADGRSDHTLTSP 2397
                        :  | ::    :  | :| : |  |   :|     |   | | |
Db       3024 --------------QQQQQQQQQHSAVLALSP-SQSPRLLTKLPGQLLPG---HGLQPP 3064

Qy       2398 GGGGKAKVSGRPSSRKAKSPAPGLASGDRPPSVSSVHSEGDCNRRTPLTNRVWEDRPSSA 2457
                |                  | | | |
Db       3065 QG-----------------PPGGQAGG-------------------------------- 3074

Qy       2458 GSTPFPYNPLIMRLQAGVMASPPPPGLP---------------AGSGPLAGP 2494
                       :||  | ||| |  || |                 |:| ||||
Db       3075 -----------LRLTPGGMALPGQPGGPFLNTALAQQQQQQHSGGAGSLAGP 3115


RESULT 16
US-10-051-874-165
; Sequence 165, Application US/10051874
; Publication No. US20040005557A1
; GENERAL INFORMATION:
;  APPLICANT: Padigaru, Muralidhara
;  APPLICANT:  Alsobrook II, John P
;  APPLICANT:  Colman, Steven D
;  APPLICANT:  Spytek, Kimberly A
;  APPLICANT:  Boldog, Ferenc
;  APPLICANT:  Vernet, Corine AM
;  APPLICANT:  Li, Li
;  APPLICANT:  Shenoy, Suresh G
;  APPLICANT:  Casman, Stacie J
;  APPLICANT:  Guo, Xiaojia Sasha
;  APPLICANT:  Edinger, Shlomit R
;  APPLICANT:  MacDougall, John R
;  APPLICANT:  Malyankar, Uriel M
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Shimkets, Richard A
;  APPLICANT:  Pena, Carol EA
;  APPLICANT:  Tchernev, Velizar T
;  APPLICANT:  Zerhusen, Bryan D
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  Miller, Charles E
;  APPLICANT:  Lepley, Denise M
;  APPLICANT:  Smithson, Glennda
;  APPLICANT:  Baumgartner, Jason C
;  APPLICANT:  Herrman, John L
;  APPLICANT:  Peyman, John A
;  APPLICANT:  Gorman, Linda
;  APPLICANT:  Mezes, Peter D
;  APPLICANT:  Kekuda, Ramesh
;  APPLICANT:  Taupier Jr, Raymond J
;  APPLICANT:  Gerlach, Valerie
;  APPLICANT:  Grosse, William M
;  APPLICANT:  Liu, Xiaohong
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  Rothenberg, Mark
;  APPLICANT:  Stone, David J
;  APPLICANT:  Burgess, Catherine E
;  TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
;  TITLE OF INVENTION:  USING THE SAME
;  FILE REFERENCE: 21402-245
;  CURRENT APPLICATION NUMBER: US/10/051,874
;  CURRENT FILING DATE:  2002-09-25
;  PRIOR APPLICATION NUMBER: 60/268,595
;  PRIOR FILING DATE: 2001-02-14
;  PRIOR APPLICATION NUMBER: 60/325,306
;  PRIOR FILING DATE: 2001-09-27
;  PRIOR APPLICATION NUMBER: 60/262,587
;  PRIOR FILING DATE: 2001-01-18
;  PRIOR APPLICATION NUMBER: 60/272,409
;  PRIOR FILING DATE: 2001-02-28
;  PRIOR APPLICATION NUMBER: 60/262,454
;  PRIOR FILING DATE: 2001-01-18
;  PRIOR APPLICATION NUMBER: 60/276,777
;  PRIOR FILING DATE: 2001-03-16
;  PRIOR APPLICATION NUMBER: 60/291,672
;  PRIOR FILING DATE: 2001-05-17
;  PRIOR APPLICATION NUMBER: 60/330,336
;  PRIOR FILING DATE: 2001-10-18
;  PRIOR APPLICATION NUMBER: 60/265,530
;  PRIOR FILING DATE: 2001-01-31
;  PRIOR APPLICATION NUMBER: 60/261,376
;  PRIOR FILING DATE: 2001-01-16
;  NUMBER OF SEQ ID NOS: 269
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 165
;   LENGTH: 5262
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-051-874-165

  Query Match             3.6%;  Score 470.5;  DB 15;  Length 5262;
  Best Local Similarity   20.3%;  Pred. No. 1.4e-12;
  Matches  558;  Conservative  284;  Mismatches  891;  Indels 1019;  Gaps  144;

Qy        145 LEPVSPPSPPHTDPELELVPPRLSKEE----------LIQNMDRVDREITMVEQQISKLK 194
              ::||:| :||      ||||  :: : |          | :   : | :||     | |
Db       1286 VKPVAPVAPP------ELVPMKVKEPEPQYFRFEGVWLTETGMALLRNLTM-----SPLH 1334

Qy        195 KKQQQ-----LEEEAA-KPPEPEKPVSPPPIESKHRSLVQIIYDENRKKAEAAHRILEG- 247
              |::|:     |  ||  :  ||   : |   :        :::  |   :       |||
Db       1335 KRRQRRGRLGLPGEAGLEGSEPSDALGPDDKKDGDLDTDELLKGEGGVEHMECEIKLEGP 1394

Qy        248 LGPQVELPLYNQPSDTRQYHENIKINQAMRKKL--------ILYFKRRNHARKQWKQKFC 299
              : ||| |   :  ::::           ||:          :  :|::| | : |
Db       1395 VSPDVE-PGKEETEESKK----------RKRKPYRPGIGGFMVRQRKSHTRTK-KGPAA 1441

Qy        300 QR----------------------YDQLMEALEKKVERIENNPRRRAKESKVREYYEKQ 336
              |                      | |  ||| ::     ||  | ||:   :
Db       1442 QAEVLSGDGQPDEVIPADLPAEGAVEQSLAEGDEKKKQQ-----RRGRKRSKLEGMFPAY 1496

Qy        337 FPEIRKQRELQERMQSR-----VGQRGSGLSMSAARSEHEVSEIIDGLSEQENLEKQMRQ 391
                |   :|| :  :       ||:   ||    |:         || ||:: |     |
Db       1497 LQEAFFGKELLDLSRKALFAVGVGRPSFGLGTPKAKG--------DGGSERKELPTS--Q 1546

Qy        392 LAVIPPMLYDADQQRIK------FINMNGLMADPMKVYKDRQVMNMWSEQEKETFREKFM 445
                | : | : ::          |: | |:             |: ||
Db       1547 KGDDGPDIADEESRGLEGKADTPGPEDGGVKASPVP----------SDPEK-------- 1587

Qy        446 QHPKNFGLIASFLERKTVAECVLYYYLTKKNENYKSLVRRSYRRRGKSQQQQQQQQQQQQ 505
                |:::| |:| :  |
Db       1588 PGTPGEGMLSSDLDRISTEE---------------------------------------- 1607

Qy        506 QQQQQPMPRSSQEEKDEKEKEKE---AEKEEE----KPEVENDKEDLLK---------- 547
                :|:   | || ::  |:   :|:|:      | :|  :  | :
Db       1608 ------LPK--MESKDLQQLFKDVLGSEREQHLGCGTPGLEGSRTPLQRPFLQGGLPLGN 1659

Qy        548 ----EKTDDTSG------EDNDEKEAVASKGRKTANSQGRRKGRITRSMANEANSEEAIT 597
                    |    |      |: |:    |      :|     | | |              |
Db       1660 LPSSSPMDSYPGLCQSPFLDSRERGGFFSPEPGEPDSPWTGSGGTTPSTPT--------T 1711

Qy        598 PQQSAELASMELNESS--RWTEEE----METAK---------KGLLEHGRNWSAIARMVG 642
              |    |   :  |: |  || ::|    : |           | :   :||
Db       1712 PTTEGEGDGLSYNQRSLQRWEKDEELGQLSTISPVLYANINFPNLKQDYPDWS------- 1764

Qy        643 SKTVSQCKNFYFNYKKRQNLDEILQQHKLKMEKERNARRKKKKAPAAASEEAAFPPVVED 702
                |:||     ::|     |:     | |     | | |   |     | |:         :
Db       1765 ----SRCKQIMKLWRKVPAADKAPYLQKAK--DNRAAHRINKVQKQAESQ--------IN 1810

Qy        703 EEMEASGVSGNEEEMVEEAEALHASGNEVPRGECSGPATVNNSSDTESIPSPHTEAAKDT 762
              :: :   ::   :        |||    |     | |     :: |  | || | |   |
Db       1811 KQTKVGDIARKTDR-----PALHLRIPPQPGALGSPPPA---AAPTIFIGSPTTPAGLST 1862

Qy        763 GQNG-PKPPATLGADGPPPGPPT--------PPRRTSRA-----------PIEP----- 796
              :|  ||||      |  ||| :         ||:  ::|           |:||
Db       1863 SADGFLKPPA-----GSVPGPDSPGELFLKLPPQVPAQAPSQDPFGLAPAYPLEPRFPTA 1917

Qy        797 ----TPASEATGAPTPPPAPPSPSAP----------PPVVPKEEKEEETAAAPPVEEGE 841
                    |    ||||  ||   : | |           ||  |: :        |
Db       1918 PPTYPPYPSPTGAPAQPPMLGASSRPGAGQPGEFHTTPPGTPRHQPSTPDPFLKP----- 1972
```

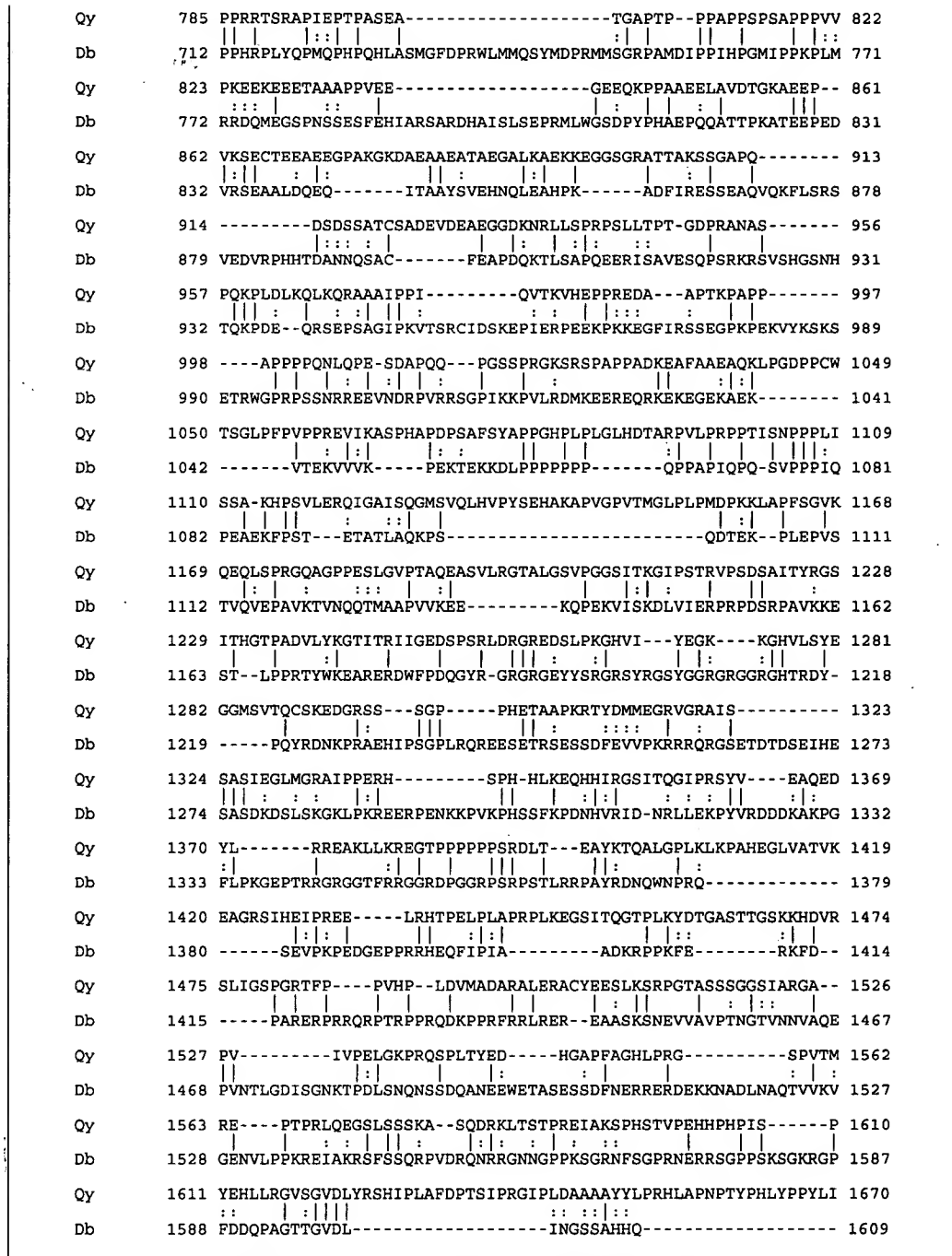


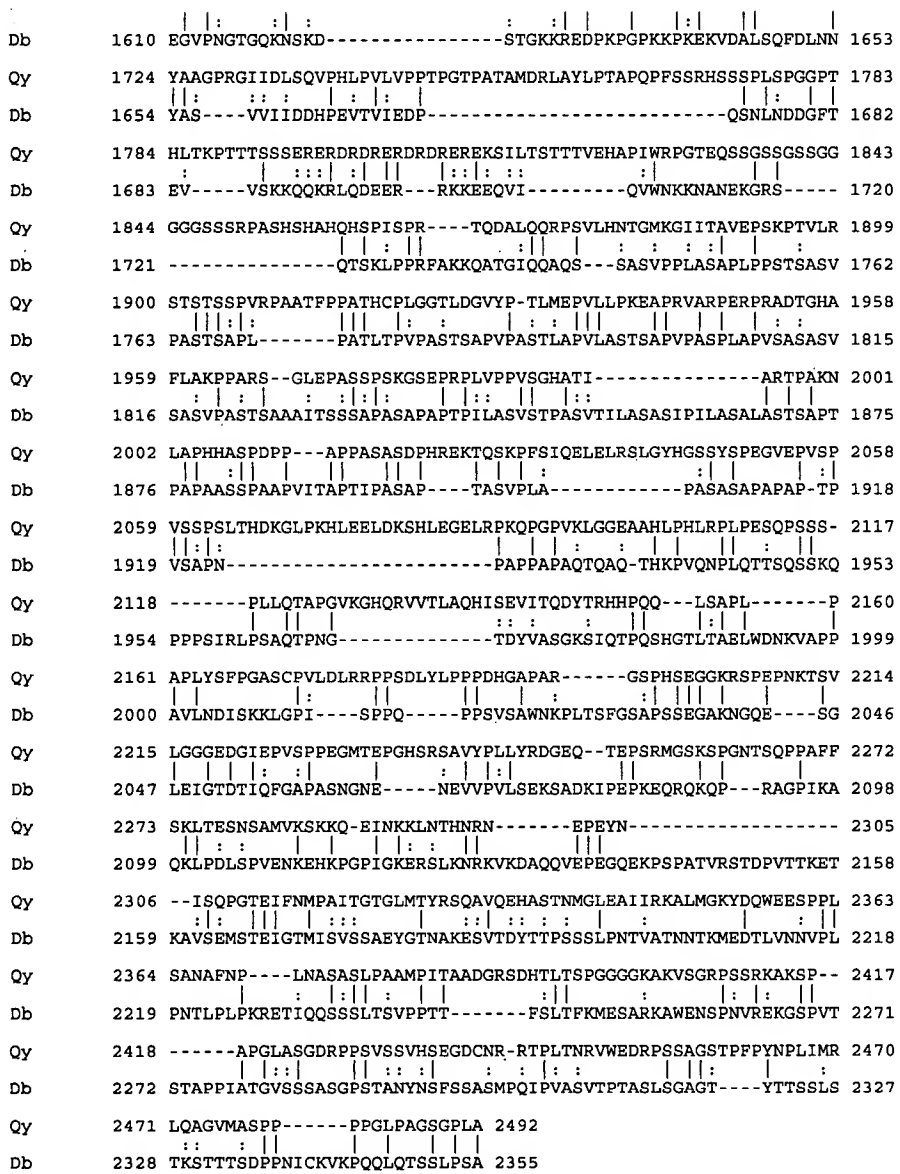


RESULT 17
US-10-171-311-83
; Sequence 83, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatt, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoersh, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER

```
;  FILE REFERENCE: MRI-035
;  CURRENT APPLICATION NUMBER: US/10/171,311
;  CURRENT FILING DATE:  2002-06-12
;  PRIOR APPLICATION NUMBER: US 60/298,159
;  PRIOR FILING DATE: 2001-06-13
;  PRIOR APPLICATION NUMBER: US 60/298,155
;  PRIOR FILING DATE: 2001-06-13
;  PRIOR APPLICATION NUMBER: US 60/335,936
;  PRIOR FILING DATE: 2001-11-14
;  NUMBER OF SEQ ID NOS: 238
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 83
;   LENGTH: 2701
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-171-311-83

  Query Match             3.6%;  Score 469.5;  DB 14;  Length 2701;
  Best Local Similarity   19.1%;  Pred. No. 6.9e-13;
  Matches  532;  Conservative  332;  Mismatches  980;  Indels  944;  Gaps  130;
```

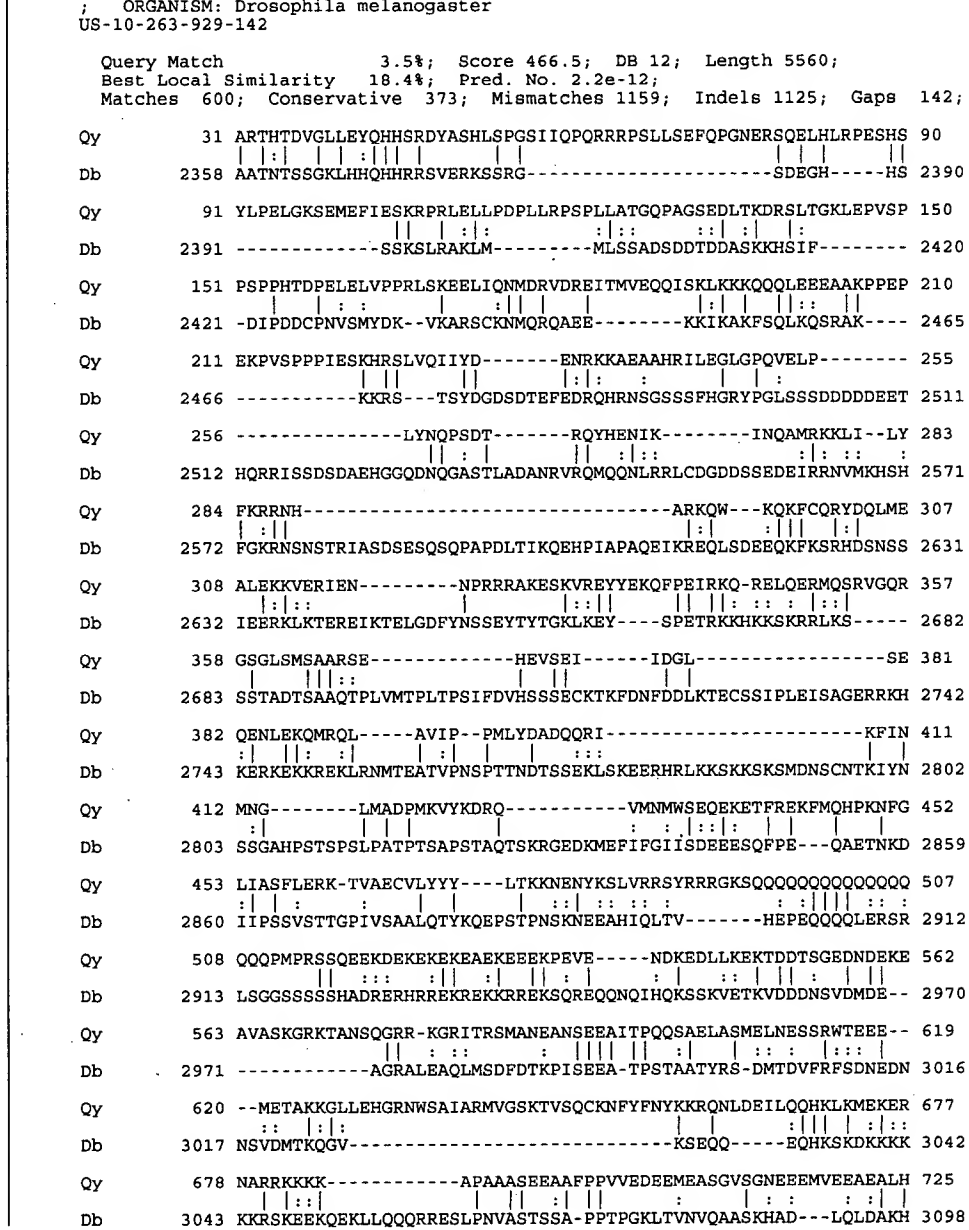


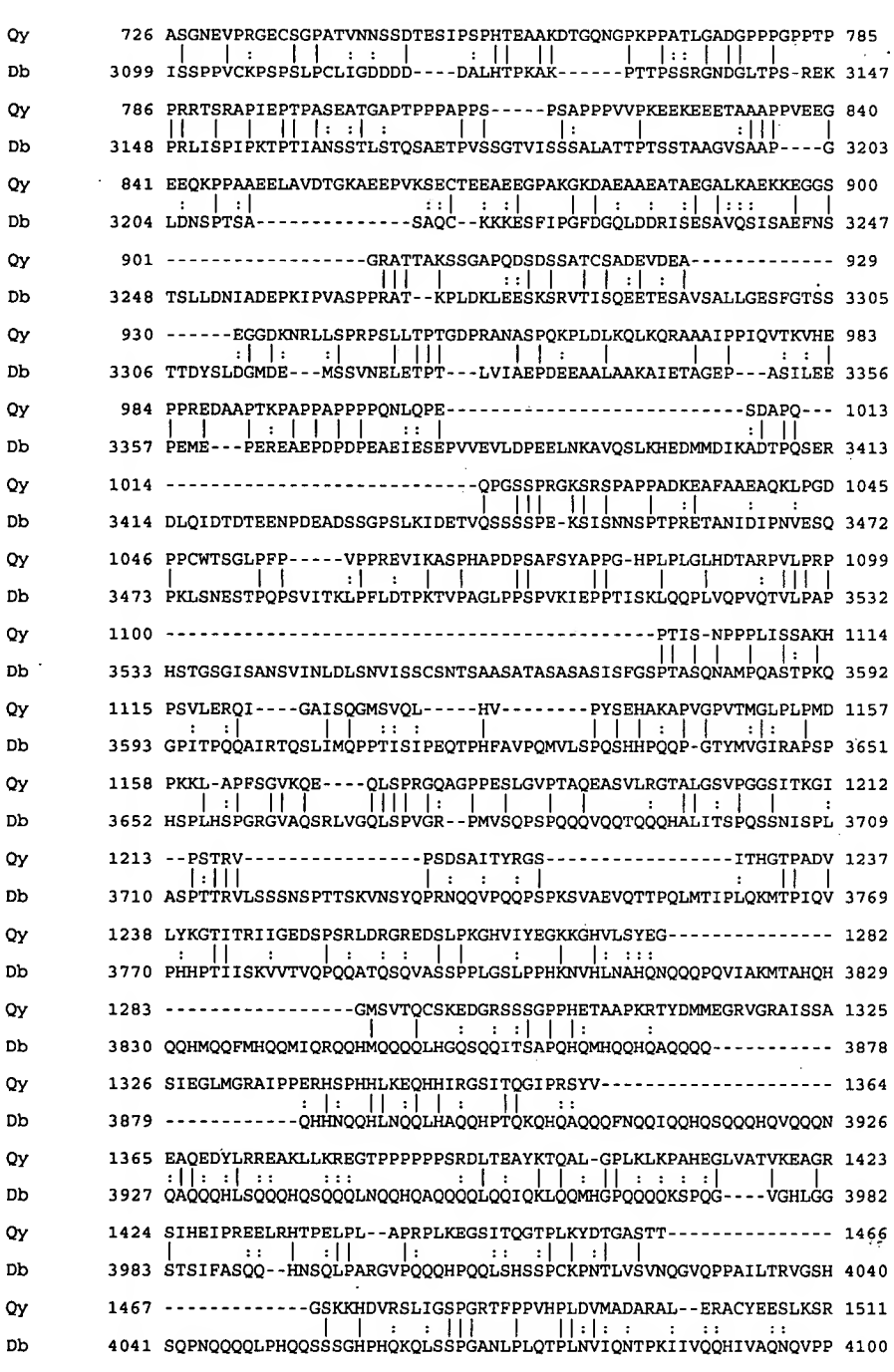


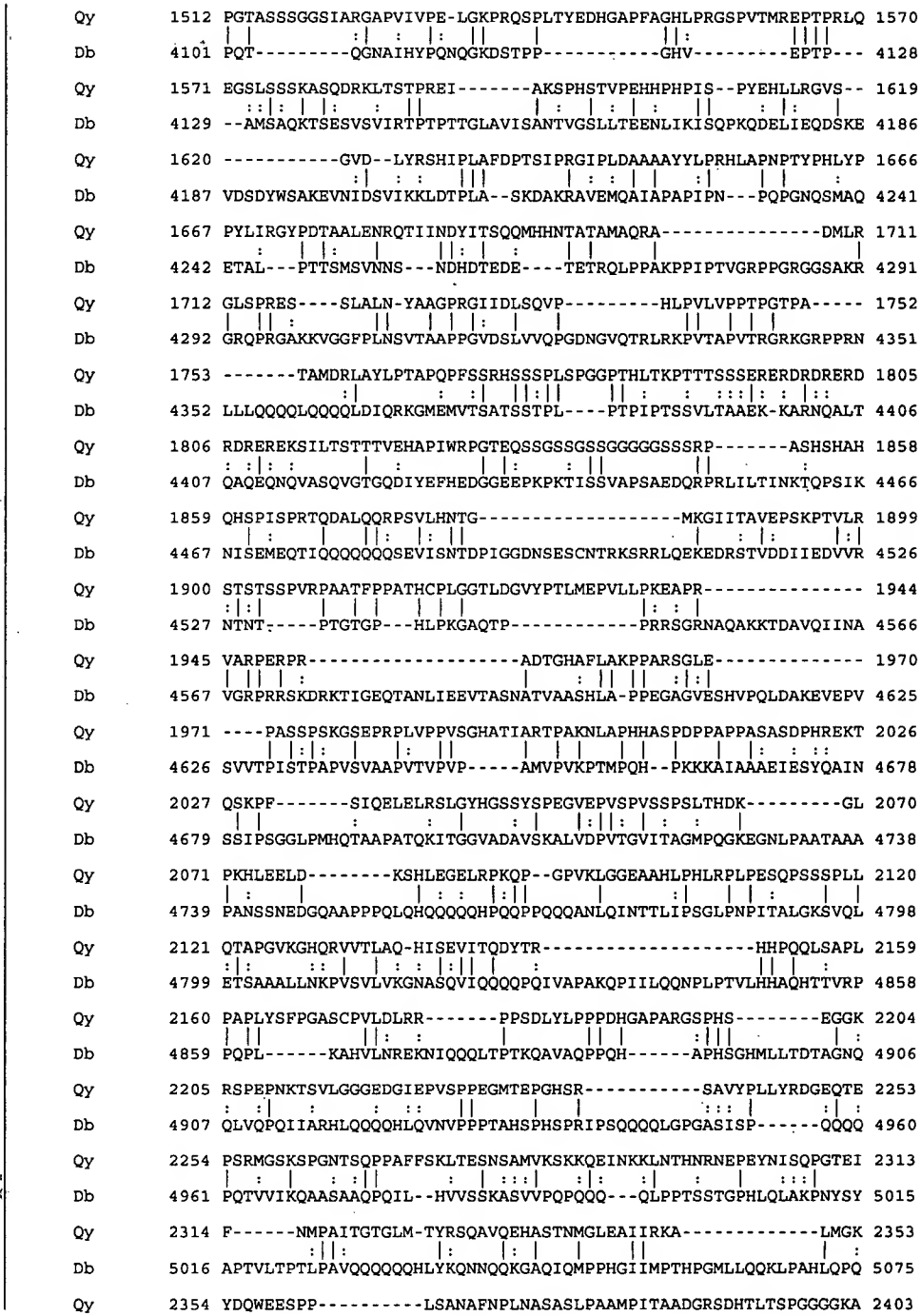


RESULT 18
US-10-263-929-142
; Sequence 142, Application US/10263929
; Publication No. US20040067535A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jaeseob
; APPLICANT: Galant, Ron
; TITLE OF INVENTION: Alzheimer's Disease Linked Genes
; FILE REFERENCE: LSD-07417
; CURRENT APPLICATION NUMBER: US/10/263,929
; CURRENT FILING DATE: 2002-10-03
; NUMBER OF SEQ ID NOS: 213
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 142
; LENGTH: 5560
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-263-929-142

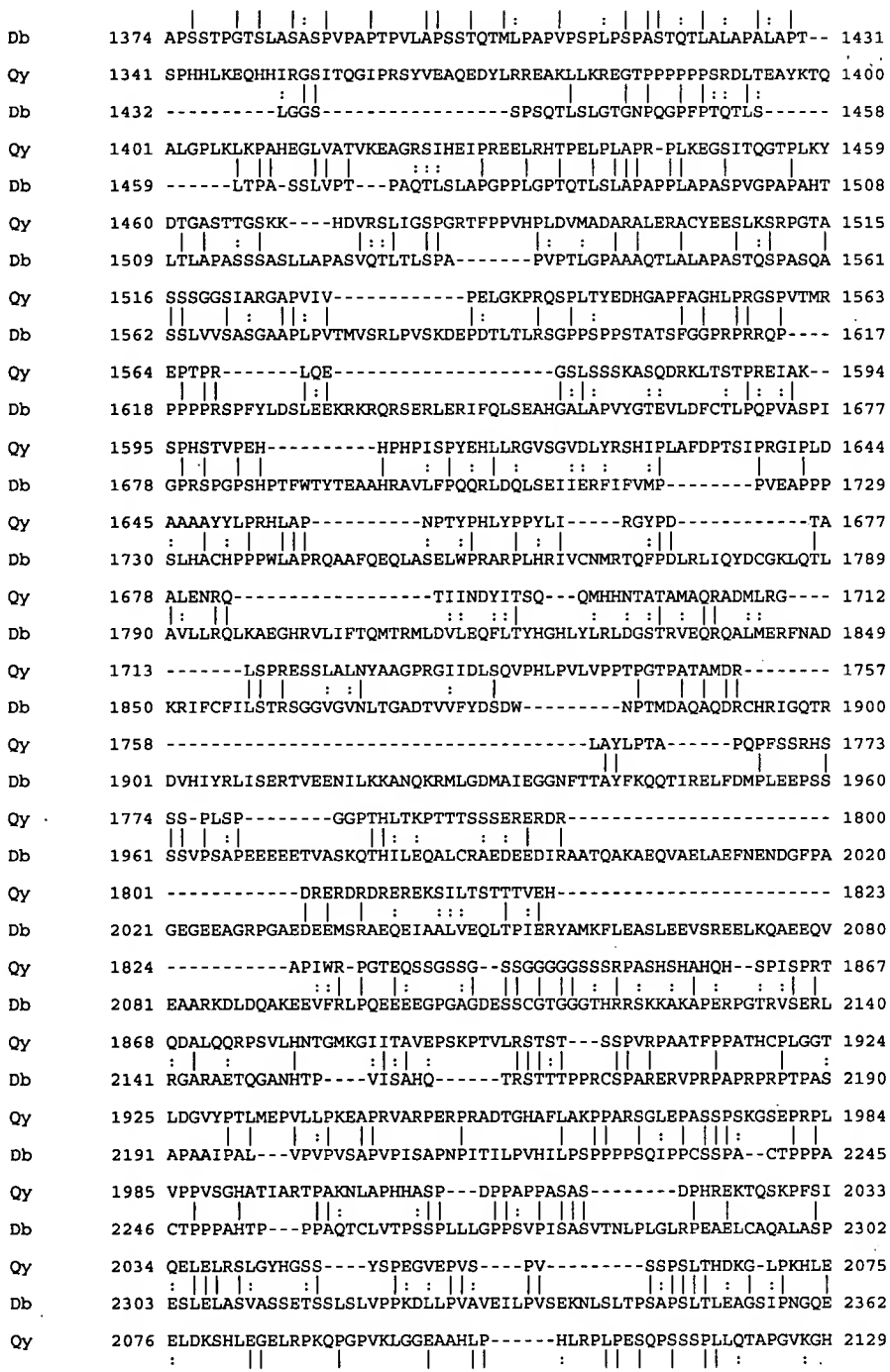
Query Match 3.5%; Score 466.5; DB 12; Length 5560;
Best Local Similarity 18.4%; Pred. No. 2.2e-12;
Matches 600; Conservative 373; Mismatches 1159; Indels 1125; Gaps 142;




Qy 726 ASGNEVPRGECSGPATVNNSSDTESIPSPHTEAAKDTGQNGPKPPATLGADGPPPGPPTP 785
Db 3099 ISSPPVCKPSPSLPCLIGDDDD----DALHTPKAK------PTTPSSRGNDGLTPS-REK 3147
Qy 786 PRRTSRAPIEPTPASEATGAPTPPPAPPS-----PSAPPPVVPKEEKEEETAAAPPVEEG 840
Db 3148 PRLISPIPKTPTIANSSTLSTQSAETPVSSGTVISSSALATTPTSSTAAGVSAAP----G 3203
Qy 841 EEQKPPAAEELAVDTGKAEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGS 900
Db 3204 LDNSPTSA--------------SAQC--KKKESFIPGFDGQLDDRISESAVQSISAEFNS 3247
Qy 901 ----------GRATTAKSSGAPQDSDSSATCSADEVDEA---------- 929
Db 3248 TSLLDNIADEPKIPVASPPRAT--KPLDKLEESKSRVTISQEETESAVSALLGESFGTSS 3305
Qy 930 ------EGGDKNRLLSPRPSLLTPTGDPRANASPQKPLDLKQLKQRAAAIPPIQVTKVHE 983
Db 3306 TTDYSLDGMDE---MSSVNELETPT---LVIAEPDEEAALAAKAIETAGEP---ASILEE 3356
Qy 984 PPREDAAPTKPAPPAPPPPQNLQPE----------SDAPQ--- 1013
Db 3357 PEME---PEREAEPDPDPEAEIESEPVVEVLDPEELNKAVQSLKHEDMMDIKADTPQSER 3413
Qy 1014 ----------QPGSSPRGKSRSPAPPADKEAFAAEAQKLPGD 1045
Db 3414 DLQIDTDTEENPDEADSSGPSLKIDETVQSSSSPE-KSISNNSPTPRETANIDIPNVESQ 3472
Qy 1046 PPCWTSGLPFP-----VPPREVIKASPHAPDPSAFSYAPPG-HPLPLGLHDTARPVLPRP 1099
Db 3473 PKLSNESTPQPSVITKLPFLDTPKTVPAGLPPSPVKIEPPTISKLQQPLVQPVQTVLPAP 3532
Qy 1100 ----------PTIS-NPPPLISSAKH 1114
Db 3533 HSTGSGISANSVINLDLSNVISSCSNTSAASATASASASISFGSPTASQNAMPQASTPKQ 3592
Qy 1115 PSVLERQI----GAISQGMSVQL-----HV--------PYSEHAKAPVGPVTMGLPLPMD 1157
Db 3593 GPITPQQAIRTQSLIMQPPTISIPEQTPHFAVPQMVLSPQSHHPQQP-GTYMVGIRAPSP 3651
Qy 1158 PKKL-APFSGVKQE----QLSPRGQAGPPESLGVPTAQEASVLRGTALGSVPGGSITKGI 1212
Db 3652 HSPLHSPGRGVAQSRLVGQLSPVGR--PMVSQPSPQQQVQQTQQQHALITSPQSSNISPL 3709
Qy 1213 --PSTRV----------PSDSAITYRGS----------ITHGTPADV 1237
Db 3710 ASPTTRVLSSSNSPTTSKVNSYQPRNQQVPQQPSPKSVAEVQTTPQLMTIPLQKMTPIQV 3769
Qy 1238 LYKGTITRIIGEDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYEG---------- 1282
Db 3770 PHHPTIISKVVTVQPQQATQSQVASSPPLGSLPPHKNVHLNAHQNQQQPQVIAKMTAHQH 3829
Qy 1283 ----------GMSVTQCSKEDGRSSSGPPHETAAPKRTYDMMEGRVGRAISSA 1325
Db 3830 QQHMQQFMHQQMIQRQQHMQQQQLHGQSQQITSAPQHQMHQQHQAQQQQ---------- 3878
Qy 1326 SIEGLMGRAIPPERHSPHHLKEQHHIRGSITQGIPRSYV---------- 1364
Db 3879 ----------QHHNQQHLNQQLHAQQHPTQKQHQAQQQFNQQIQQHQSQQQHQVQQQN 3926
Qy 1365 EAQEDYLRREAKLLKREGTPPPPPPSRDLTEAYKTQAL-GPLKLKPAHEGLVATVKEAGR 1423
Db 3927 QAQQQHLSQQQHQSQQQLNQQHQAQQQQLQQIQKLQQMHGPQQQQKSPQG----VGHLGG 3982
Qy 1424 SIHEIPREELRHTPELPL--APRPLKEGSITQGTPLKYDTGASTT---------- 1466
Db 3983 STSIFASQQ--HNSQLPARGVPQQQHPQQLSHSSPCKPNTLVSVNQGVQPPAILTRVGSH 4040
Qy 1467 ----------GSKKHDVRSLIGSPGRTFPPVHPLDVMADARAL--ERACYEESLKSR 1511
Db 4041 SQPNQQQQLPHQQSSSGHPHQKQLSSPGANLPLQTPLNVIQNTPKIIVQQHIVAQNQVPP 4100

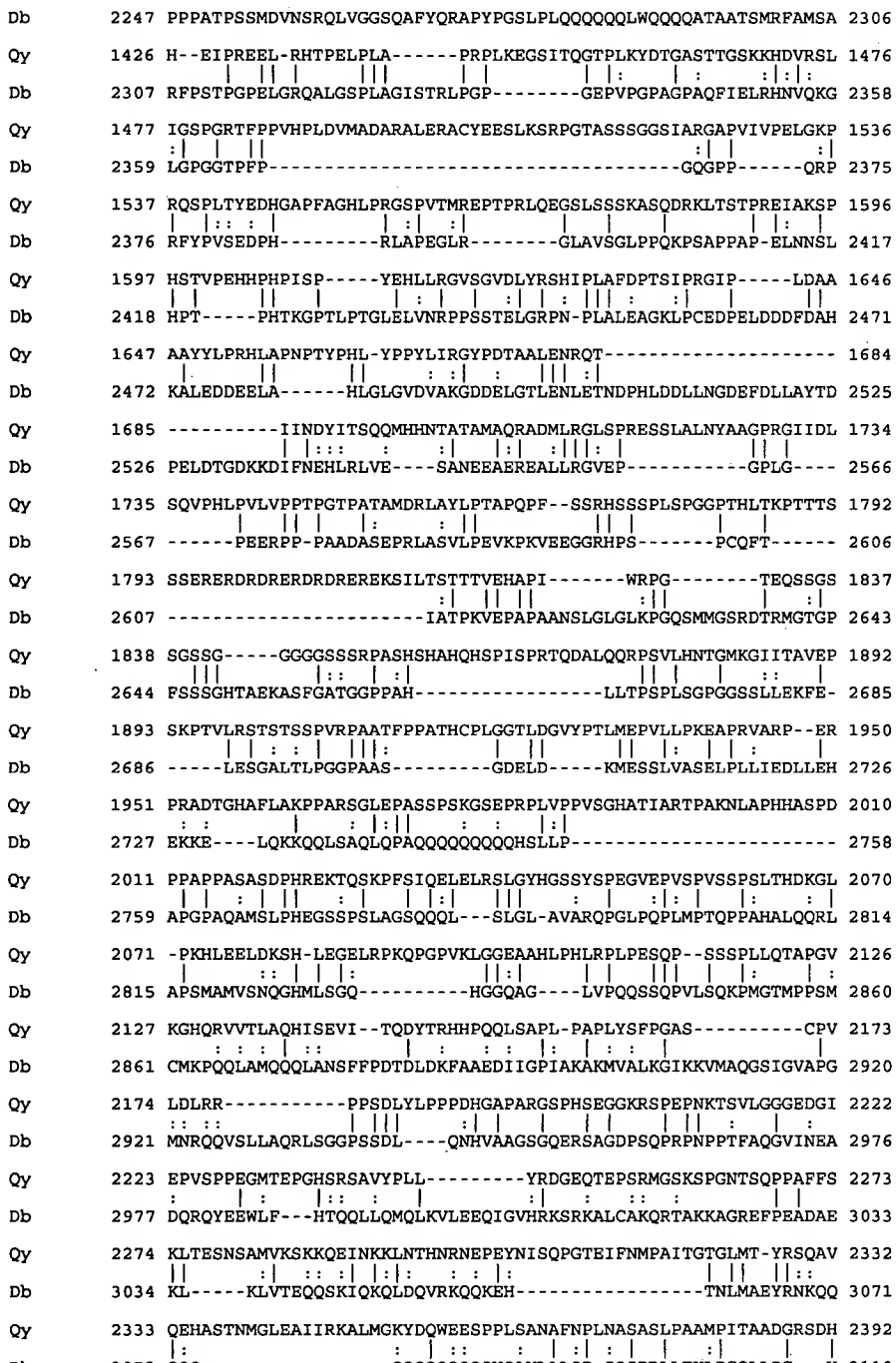

Qy 1512 PGTASSSGGSIARGAPVIVPE-LGKPRQSPLTYEDHGAPFAGHLPRGSPVTMREPTPRLQ 1570
Db 4101 PQT---------QGNAIHYPQNQGKDSTPP----------GHV---------EPTP--- 4128
Qy 1571 EGSLSSSKASQDRKLTSTPREI-------AKSPHSTVPEHHPHPIS--PYEHLLRGVS-- 1619
Db 4129 --AMSAQKTSESVSVIRTPTPTTGLAVISANTVGSLLTEENLIKISQPKQDELIEQDSKE 4186
Qy 1620 ----------GVD--LYRSHIPLAFDPTSIPRGIPLDAAAAYYLPRHLAPNPTYPHLYP 1666
Db 4187 VDSDYWSAKEVNIDSVIKKLDTPLA--SKDAKRAVEMQAIAPAPIPN---PQPGNQSMAQ 4241
Qy 1667 PYLIRGYPDTAALENRQTIINDYITSQQMHHNTATAMAQRA----------DMLR 1711
Db 4242 ETAL---PTTSMSVNNS---NDHDTEDE----TETRQLPPAKPPIPTVGRPPGRGGSAKR 4291
Qy 1712 GLSPRES----SLALN-YAAGPRGIIDLSQVP---------HLPVLVPPTPGTPA----- 1752
Db 4292 GRQPRGAKKVGGFPLNSVTAAPPGVDSLVVQPGDNGVQTRLRKPVTAPVTRGRKGRPPRN 4351
Qy 1753 -------TAMDRLAYLPTAPQPFSSRHSSSPLSPGGPTHLTKPTTTSSSERERDRDRERD 1805
Db 4352 LLLQQQQLQQQQLDIQRKGMEMVTSATSSTPL----PTPIPTSSVLTAAEK-KARNQALT 4406
Qy 1806 RDREREKSILTSTTTVEHAPIWRPGTEQSSGSSGSSGGGGGSSSRP-------ASHSHAH 1858
Db 4407 QAQEQNQVASQVGTGQDIYEFHEDGGEEPKPKTISSVAPSAEDQRPRLILTINKTQPSIK 4466
Qy 1859 QHSPISPRTQDALQQRPSVLHNTG----------MKGIITAVEPSKPTVLR 1899
Db 4467 NISEMEQTIQQQQQQQSEVISNTDPIGGDNSESCNTRKSRRLQEKEDRSTVDDIIEDVVR 4526
Qy 1900 STSTSSPVRPAATFPPATHCPLGGTLDGVYPTLMEPVLLPKEAPR---------- 1944
Db 4527 NTNT-----PTGTGP---HLPKGAQTP----------PRRSGRNAQAKKTDAVQIINA 4566
Qy 1945 VARPERPR----------ADTGHAFLAKPPARSGLE---------- 1970
Db 4567 VGRPRRSKDRKTIGEQTANLIEEVTASNATVAASHLA-PPEGAGVESHVPQLDAKEVEPV 4625
Qy 1971 ----PASSPSKGSEPRPLVPPVSGHATIARTPAKNLAPHHASPDPPAPPASASDPHREKT 2026
Db 4626 SVVTPISTPAPVSVAAPVTVPVP-----AMVPVKPTMPQH--PKKKAIAAAEIESYQAIN 4678
Qy 2027 QSKPF-------SIQELELRSLGYHGSSYSPEGVEPVSPVSSPSLTHDK---------GL 2070
Db 4679 SSIPSGGLPMHQTAAPATQKITGGVADAVSKALVDPVTGVITAGMPQGKEGNLPAATAAA 4738
Qy 2071 PKHLEELD--------KSHLEGELRPKQP--GPVKLGGEAAHLPHLRPLPESQPSSSPLL 2120
Db 4739 PANSSNEDGQAAPPPQLQHQQQQQHPQQPPQQQANLQINTTLIPSGLPNPITALGKSVQL 4798
Qy 2121 QTAPGVKGHQRVVTLAQ-HISEVITQDYTR----------HHPQQLSAPL 2159
Db 4799 ETSAAALLNKPVSVLVKGNASQVIQQQQPQIVAPAKQPIILQQNPLPTVLHHAQHTTVRP 4858
Qy 2160 PAPLYSFPGASCPVLDLRR-------PPSDLYLPPPDHGAPARGSPHS--------EGGK 2204
Db 4859 PQPL------KAHVLNREKNIQQQLTPTKQAVAQPPQH------APHSGHMLLTDTAGNQ 4906
Qy 2205 RSPEPNKTSVLGGGEDGIEPVSPPEGMTEPGHSR----------SAVYPLLYRDGEQTE 2253
Db 4907 QLVQPQIIARHLQQQQHLQVNVPPPTAHSPHSPRIPSQQQQLGPGASISP------QQQQ 4960
Qy 2254 PSRMGSKSPGNTSQPPAFFSKLTESNSAMVKSKKQEINKKLNTHNRNEPEYNISQPGTEI 2313
Db 4961 PQTVVIKQAASAAQPQIL--HVVSSKASVVPQPQQQ---QLPPTSSTGPHLQLAKPNYSY 5015
Qy 2314 F------NMPAITGTGLM-TYRSQAVQEHASTNMGLEAIIRKA----------LMGK 2353
Db 5016 APTVLTPTLPAVQQQQQQHLYKQNNQQKGAQIQMPPHGIIMPTHPGMLLQQKLPAHLQPQ 5075
Qy 2354 YDQWEESPP----------LSANAFNPLNASASLPAAMPITAADGRSDHTLTSPGGGGKA 2403

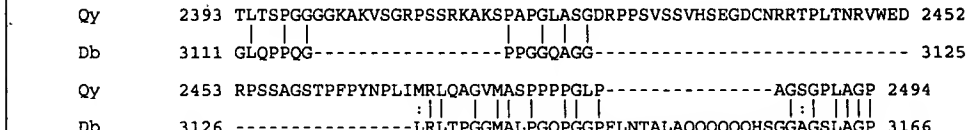
```
Db       5076 QHQLNPSPPPGKPNPVLHGLQSGQIMPGSVGSPPPVSAAVLKTAQQQVNSVVPVAGIRTA 5135

Qy       2404 KVSGRPSSRKAKSP---APGLASGDRPPSVSSVHSEGD--CNRRT----------PLTN 2447

Db       5136 IPNISPQSQPRVSPLVLPPGISG--VPPFDASLHDLGAYVSGRRTQSPPPAHQQASPITP 5193

Qy       2448 RVWEDRPSSAGSTPFPYNPLIMRLQAG----VMASPPPPGL-PAGSGP---LAGPHHAWD 2499

Db       5194 NDSTYRGVTASRDFMLYQHHLMR--GGDYDDKMGSSPPLELRRPGSGPPRTIAVPH---- 5247

Qy       2500 EEPKPLLCSQYETLSDS 2516

Db       5248 ----SLQSPQDRTAADS 5260


RESULT 19
US-10-146-473-50
; Sequence 50, Application US/10146473
; Publication No. US20030108888A1
; GENERAL INFORMATION:
;  APPLICANT: Scanlan, Matthew
;  APPLICANT:  Gout, Ivan
;  APPLICANT:  Stockert, Elisabeth
;  APPLICANT:  Gure, Ali
;  APPLICANT:  Chen, Yao-Tseng
;  APPLICANT:  Old, Lloyd
;  TITLE OF INVENTION: Breast Cancer Antigens
;  FILE REFERENCE: L00461/70130(JRV)
;  CURRENT APPLICATION NUMBER: US/10/146,473
;  CURRENT FILING DATE:  2002-05-15
;  PRIOR APPLICATION NUMBER: US 60/291,150
;  PRIOR FILING DATE: 2001-05-15
;  NUMBER OF SEQ ID NOS: 82
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 50
;   LENGTH: 2971
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-146-473-50

  Query Match             3.4%;  Score 452;  DB 14;  Length 2971;
  Best Local Similarity   19.0%;  Pred. No. 4.9e-12;
  Matches  587;  Conservative  296;  Mismatches 1107;  Indels 1092;  Gaps  130;

Qy         71 EFQPGNERSQE----LHLRPESHSYLPELGKSEMEFI--ESKRPRLELLPD--PLLRPSP 122

Db         91 DFQPQEDEEEDDEETIEVEEQQEGNDAEAQRREIELLRREGELPLEELLRSLPPQLLEGP 150

Qy        123 LLATGQPAGSEDLTKDRSLTGKLEPVSPPSPPHTDPELELVPPR---LSKEELIQNMDRV 179

Db        151 SSPSQTPSSHDSDTRDGPEEGAEE-----EPPQV---LEIKPPPSAVTQRNKQPWHPDED 202

Qy        180 DREITMVEQQISKLK---KKQQQLEEEAAKPPEPEKPVSPPPIESKHRSLVQIIY----- 231

Db        203 DEEFTANEEEAEDEEDTIAAEEQLEGEVDHAMELSELAREGEL-----SMEELLQQYAGA 257

Qy        232 ---------DENRKKAEAAHRILEGLGP-QVELPLYNQPSDTRQYHENIKINQAMRKKLI 281

Db        258 YAPGSGSSEDEDEDEVDANSSDCEPEGPVEAEEP---PQEDSSSQSDSVE---------- 304

Qy        282 LYFKRRNHARKQWKQKFCQRYDQLMEALEKKVERIENNPRRRAKESKVREYYEKQFPEIR 341

Db        305 ---------------------DRSEDEEDEHSEEEETSGSSASEESESEESEDAQSQSQA 343

Qy        342 KQRE----------LQERMQSRVGQRGSGLSMSAARSEHEVSEIIDGLSEQENLEKQMRQ 391

Db        344 DEEEEDDDFGVEYLLARDEEQSEADAGSGPPTPGPTTLGPKKEITDIAAAAESLQPKGYT 403

Qy        392 LAVIP-----PMLYDADQQRIKFI----------NMNGLMADPMKVYKDRQVMNM---- 431

Db        404 LATTQVKTPIPLLLRGQLREYQHIGLDWLVTMYEKKLNGILADEMGLGKTIQTISLLAHL 463

Qy        432 ------WS-------------EQEKETFREKFMQHPKNPGLIASFLERKTVAE------ 465

Db        464 ACEKGNWGPHLIIVPTSVMLNWEMELKRWCPSF----KILTYYGAQKERKLKRQGWTKPN 519

Qy        466 ----CVLYYYLT--------KKNENYKSLVRRSYRRRGKSQQQQQQQQQQQQQQ------ 507

Db        520 AFHVCITSYKLVLQDHQAFRRKNWRYLILDEAQNIKNFKSQRWQSLLNFNSQRRLLLTGT 579

Qy        508 ------------QQQPMPRSSQEEKDEKE---------KEKEAEKEEE---------KP- 536

Db        580 PLQNSLMELWSLMHFLMPHVFQSHREFKEWFSNPLTGMIEGSQEYNEGLVKRLHKVLRPF 639

Qy        537 ---EVENDKEDLLKEKTD--------------DTSGEDNDEKEAVASKGRKTANSQGRR 578

Db        640 LLRRVKVDVEKQMPKKYEHVIRCRLSKRQRCLYDDFMAQTTTKETLAT---------GHF 690

Qy        579 KGRIT-----RSMANEAN-------SEEAITPQQSAELASMELNESSRWTEEEMETAKKG 626

Db        691 MSVINILMQLRKVCNHPNLFDPRPVTSPFITPGICFSTASLVLRATDVHPLQRIDMGRFD 750

Qy        627 LLE-HGRNWSAIARMVGSKTVSQCKNFYFNYKKRQNLDEILQQHKLKMEKERNARRKKKK 685

Db        751 LIGLEGR-----------------------VSRYEADTFLPRHRL-------SRRVLLE 779

Qy        686 APAAASEEAAFPPVVEDEEMEASGVSGNEEEMVEEAEALHASGNEVPRGECSGPATVNNS 745

Db        780 VATAPDP----PPRPKPVKMKVN-------RMLQ----------PVPKQEGRTVVVVNNP 818

Qy        746 ----SDTESIPSPHTE-AAKDTGQNGPKP---PATLGADGPPPGPP-----TPPRRTSRA 792

Db        819 RAPLGPVPVRPPPGPELSAQPT--PGPVPQVLPASLMVSASPAGPPLIPASRPPGPVLLP 876

Qy        793 PIEPTPASEATGAPTP---------PPAP-----PSPSAPPPVVPKEEKEEETAAAPPVE 838

Db        877 PLQPNSGSLPQVLPSPLGVLSGTSRPPTPTLSLKPTPPAPVRLSP----------APPPG 926

Qy        839 EGEEQKPPAAEELAVDTGKAEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEG 898

Db        927 SSSLLKP-----LTVPPGYTFPP------------------------------------ 944

Qy        899 GSGRATTAKSSGAPQDSDSSATCSADEVDEAEGGDKNRLLSPRPSLLTPTGD-------- 950

Db        945 --AAATTT--------STTTATATTTAVPAPTPAPQRLILSPDMQARLPSGEVVSIGQLA 994

Qy        951 -----PRANASPQKPLDLK----------QLKQ------RAAAIPPIQVTK-------V 981

Db        995 SLAQRPVANAGGSKPLTFQIQGNKLTLTGAQVRQLAVGQPRPLQMPPTMVNNTGVVKIVV 1054

Qy        982 HEPPREDAAPTKPAPPAPPPPQNLQPESDAPQ---QPGSSPR---GKSRSPAP------- 1028

Db       1055 RQAPRDGLTPVPPLAPAPRPPSSGLPAVLNPRPTLTPGRLPTPTLGTARAPMPTPTLVRP 1114

Qy       1029 -------PADKEAFAAEAQKLPGDPPCWTSGLPFPVP---------PREVIKASPHAPDP 1072

Db       1115 LLKLVHSPSPEVSASA-----PGAAPL-TISSPLHVPSSLPGPASSPMPIPNSSPLASPV 1168

Qy       1073 SAFSYAPPGHPLPLGLHDTARPVLPRPPTISNPPPLISSAKHPSVLERQIGAISQGMSVQ 1132

Db       1169 SSTVSVPLSSSLPISVPTTLPAPASAPLTIPISAPLTVSASGPALL-------------- 1214

Qy       1133 LHVPYSEHAKAPVGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPES---LGVPTA 1189

Db       1215 ------TSVTPPLAPVVPAAPGP-------------PSLQPSG-ASPSASALTLGLATA 1253

Qy       1190 Q--EASVLRGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHG--------------- 1232

Db       1254 PSLSSSQTPGHPLLLAPTSSHVPGLNSTVAPACSPVLVPASALASPFPSAPNPAPAQASL 1313

Qy       1233 -TPADVLYKGTITRIIGEDSPSRLDRGREDSLPKGHVIY----EGKKGHVLSYEGGMSVT 1287

Db       1314 LAPASSASQALATPLAPMAAPQTAILAPSPAPPLAPLPVLAPSPGAAPVLASSQTPVPVM 1373

Qy       1288 QCSKEDGRS--SSGP---PHETAAPKRTYDMMEGRVGRAI-SSASIEGL-MGRAIPPERH 1340
```



RESULT 20
US-10-051-874-166
; Sequence 166, Application US/10051874
; Publication No. US20040005557A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Alsobrook II, John P
; APPLICANT: Colman, Steven D
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Boldog, Ferenc
; APPLICANT: Vernet, Corine AM
; APPLICANT: Li, Li
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Casman, Stacie J
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Edinger, Shlomit R
; APPLICANT: MacDougall, John R
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard A
; APPLICANT: Pena, Carol EA
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Millet, Isabelle
; APPLICANT: Miller, Charles E
; APPLICANT: Lepley, Denise M
; APPLICANT: Smithson, Glennda
; APPLICANT: Baumgartner, Jason C
; APPLICANT: Herrman, John L
; APPLICANT: Peyman, John A
; APPLICANT: Gorman, Linda
; APPLICANT: Mezes, Peter D
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Gerlach, Valerie
; APPLICANT: Grosse, William M
; APPLICANT: Liu, Xiaohong
; APPLICANT: Ellerman, Karen

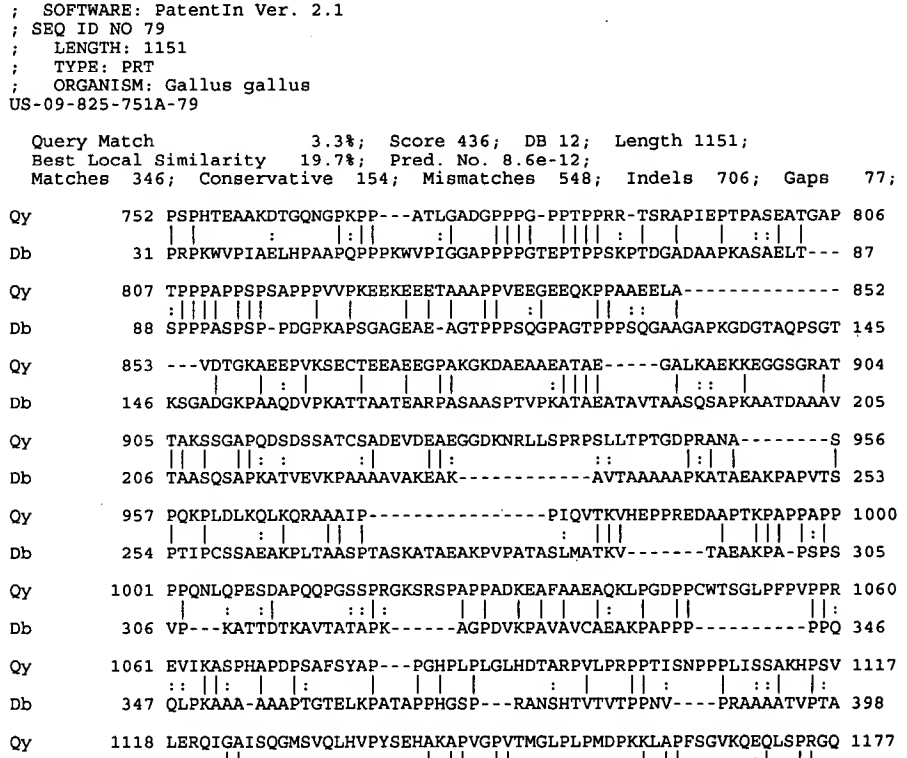
```
;  APPLICANT:  Rothenberg, Mark
;  APPLICANT:  Stone, David J
;  APPLICANT:  Burgess, Catherine E
;  TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
;  TITLE OF INVENTION:  USING THE SAME
;  FILE REFERENCE: 21402-245
;  CURRENT APPLICATION NUMBER: US/10/051,874
;  CURRENT FILING DATE:  2002-09-25
;  PRIOR APPLICATION NUMBER: 60/268,595
;  PRIOR FILING DATE: 2001-02-14
;  PRIOR APPLICATION NUMBER: 60/325,306
;  PRIOR FILING DATE: 2001-09-27
;  PRIOR APPLICATION NUMBER: 60/262,587
;  PRIOR FILING DATE: 2001-01-18
;  PRIOR APPLICATION NUMBER: 60/272,409
;  PRIOR FILING DATE: 2001-02-28
;  PRIOR APPLICATION NUMBER: 60/262,454
;  PRIOR FILING DATE: 2001-01-18
;  PRIOR APPLICATION NUMBER: 60/276,777
;  PRIOR FILING DATE: 2001-03-16
;  PRIOR APPLICATION NUMBER: 60/291,672
;  PRIOR FILING DATE: 2001-05-17
;  PRIOR APPLICATION NUMBER: 60/330,336
;  PRIOR FILING DATE: 2001-10-18
;  PRIOR APPLICATION NUMBER: 60/265,530
;  PRIOR FILING DATE: 2001-01-31
;  PRIOR APPLICATION NUMBER: 60/261,376
;  PRIOR FILING DATE: 2001-01-16
;  NUMBER OF SEQ ID NOS: 269
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 166
;   LENGTH: 5008
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-051-874-166

  Query Match             3.4%;  Score 451;  DB 15;  Length 5008;
  Best Local Similarity   20.2%;  Pred. No. 1e-11;
  Matches  569;  Conservative  290;  Mismatches  952;  Indels 1006;  Gaps  143;

Qy         53 LSPGSIIQPQRRRPSLLSEFQPGNERSQELHLRPESHSYLPELGKSEMEFIESKRPRLEL 112
              : | : : |     |  : | :|   | : : |      : |  |  | :     :| || |
Db        981 VKPVAPVAPPELVPMKVKEPEPQYFRFEGVWLTETGMALLRNLTMSPLHKRRQRRGRLGL 1040

Qy        113 LPDPLLRPS-PLLATGQPAGSE-DLTKDRSLTG-----------KLE-PVSPPSPPHTDP 158
                 :  | | |  | |     : ||  |  | |           ||| ||||   |
Db       1041 PGEAGLEGSEPSDALGPDDKKDGDLDTDELLKGEGGVEHMECEIKLEGPVSPDVEP---- 1096

Qy        159 ELELVPPRLSKEELIQNMDRVDREIT------MVEQQIS--KLKKKQQQLEEEAAKPPEP 210
                  |||  ::  |  :          || |: |  : ||       |  :   :|
Db       1097 ---------GKEETEESKKRKRKPYRPGIGGFMVRQRKSHTRTKKGPAAQAEVLSGDGQP 1147

Qy        211 EK--PVSPPPIESKHRSLVQIIYDENRKKAEAAHRILEGLGPQVE-LPLYNQPSDTRQYH 267
              ::  |   |   :  :|| :    : :|| :   |   |   | | |    || :
Db       1148 DEVIPADLPAEGAVEQSLAE---GDEKKKQQRRGRKRSGPAAQAEVLSGDGQPDEV--IP 1202

Qy        268 ENIKINQAMRKKLILYFKRRNHARKQWKQKFCQRY----------------DQLMEALE 310
               ::    |: : |    :::   |:  |: |  :                 | :|:
Db       1203 ADLPAEGAVEQSLAEGDEKKKQQRRGRKKSFFAQLAGETTLDGQPIERTIDEDNIMD--P 1260

Qy        311 KKVERIENNPRRRA-KESKVREYYEKQFPEIRKQRELQERMQSR-----VGQRGSGLSMS 364
              |  |  |   :|| |:||:   :      |    :|| :  :       ||:   ||
Db       1261 KPAEGEEQAKKRRGRKKSKLEGMFPAYLQEAFFGKELLDLSRKALFAVGVGRPSFGLGTP 1320

Qy        365 AARSEHEVSEIIDGLSEQENLEKQMRQLAVIPPMLYDADQQRIK------FINMNGLMAD 418
               |:           || ||:: |      |     | : | : : ::          |: |
Db       1321 KAKG--------DGGSERKELPTS--QKGDDGPDIADEESRGLEGKADTPGPEDGGVKAS 1370

Qy        419 PMKVYKDRQVMNMWSEQEKETFREKFMQHPKNFGLIASFLERKTVAECVLYYYLTKKNEN 478
              |:           |: ||          |:::| |:| :  |
Db       1371 PVP----------SDPEK--------PGTPGEGMLSSDLDRISTEE------------- 1398

Qy        479 YKSLVRRSYRRRGKSQQQQQQQQQQQQQQQQQQPMPRSSQEEKDEKEKEKE---AEKEEE- 534
                                                :|:   | || ::  |:   :|:|:
Db       1399 --------------------------------LPK--MESKDLQQLFKDVLGSEREQHL 1423

Qy        535 ---KPEVENDKEDLLK--------------EKTDDTSG------EDNDEKEAVASKGRK 570
                  | :|  :  | :                  |   |        |: |:    |
Db       1424 GCGTPGLEGSRTPLQRPFLQGGLPLGNLPSSSPMDSYPGLCQSPFLDSRERGGFFSPEPG 1483

Qy        571 TANSQGRRKGRITRSMANEANSEEAITPQQSAELASMELNESS--RWTEEE----METAK 624
               :|      |  | |            ||     |   :  |: |  || ::|    : |
Db       1484 EPDSPWTGSGGTTPSTPT--------TPTTEGEGDGLSYNQRSLQRWEKDEELGQLSTIS 1535

Qy        625 ---------KGLLEHGRNWSAIARMVGSKTVSQCKNFYFNYKKRQNLDEILQQHKLKMEK 675
                        | :   :||           |:||     ::|     |:      | |
Db       1536 PVLYANINFPNLKQDYPDWS----------SRCKQIMKLWRKVPAADKAPYLQKAK--D 1582

Qy        676 ERNARRKKKKAPAAASEEAAFPPVVEDEEMEASGVSGNEEEMVEEAEALHASGNEVPRGE 735
              | | |  |    | |:          ::: :   ::   :        |||      |
Db       1583 NRAAHRINKVQKQAESQ--------INKQTKVGDIARKTDR-----PALHLRIPPQPGAL 1629

Qy        736 CSGPATVNNSSDTESIPSPHTEAAKDTGQNG-PKPPATLGADGPPPGPPT--------PP 786
              | |      :: |  | || | |    |  :|  ||||      |  ||| :          ||
Db       1630 GSPPPA---AAPTIFIGSPTTPAGLSTSADGFLKPPA-----GSVPGPDSPGELFLKLPP 1681

Qy        787 RRTSRA------------PIEP---------TPASEATGAPTPPPAPPSPSAP------- 818
               :  ::|           |:||         |    ||||  ||   : | |
Db       1682 QVPAQAPSQDPFGLAPAYPLEPRFPTAPPTYPPYPSPTGAPAQPPMLGASSRPGAGQPGE 1741

Qy        819 ----PPVVPKEEKEEETAAAPPVEEGEEQKPPAAEELAVDT------GKAEEPVKSECTE 868
                  ||  |: :           |       : |: : |||         ||| ||: |
Db       1742 FHTTPPGTPRHQPSTPDPFLKP-------RCPSLDNLAVPESPGVGGGKASEPLLSP--- 1791

Qy        869 EAEEGPAKGKDAEAAEATAEGALKAEKKEGGSGRATTAKSSGAPQ----DSDSSAT---- 920
                  |  |:  :          ||: :|:| |:    :: | | |     || || |
Db       1792 -----PPFGESRK--------ALEVKKEELGA----SSPSYGPPNLGFVDSPSSGTHLGG 1834

Qy        921 --CSADEVDEAEGGDKNRLLSPRPSLLTPTGDPRANASPQKPLDLKQLKQRAAAIPPIQV 978
                   :| :|       |:|| | : |   |     ||
Db       1835 LELKTPDVFKAP-------LTPRASQVEPQ-SPGLGLRPQ-------------------- 1866

Qy        979 TKVHEPPREDAAPTKPAPPAPPPPQNLQPESDAPQQPGSSPRGKSRSPAPPADKEAFAAE 1038
                 |||   : |: |     |  ||: |  | :|  |          |  |||
Db       1867 ----EPPPAQSLPSDPFSRVPVSPQS-QSSSQSPLTP-----------RPLSAEAFC-- 1907

Qy       1039 AQKLPGDPPCWTSGLPFPVPPREVIKASPHAPDPSAFSYAPPGHPL---PLG-LHDTARP 1094
                          | || ||        :||| :    ||  |    |   ||   :|
Db       1908 --------------PSPVTPRF------QSPDPYS---RPPSRPQSRDPFAPLH---KP 1940

Qy       1095 VLPRPPTISNPPPLISSAKHPSVLERQIGAISQGMSVQLHV-PYSE-HAKAPVG------ 1146
                |:|| :       : |  |:    :||   |    |   |  | ||| | |
Db       1941 PRPQPPEV--------AFKAGSLAHTSLGA--GGFPAALPAGPAGELHAKVPSGQPPNFV 1990

Qy       1147 --PVT---MGLPLPM---------DPKKLAPFSGVKQEQLSP----RGQAGPPESLGVPT 1188
                | |   :| | ||          :|    |   | |  | |          ||  :|| |
Db       1991 RSPGTGAFVGTPSPMRFTFPQAVGEPSLKPP---VPQPGLPPPHGINSHFGPGPTLGKPQ 2047

Qy       1189 AQEASVL------RGTALGSVPGGS-ITKGIPSTRVPS-------DSAITYRGSITHGTP 1234
                :   :|        |: ||   | :  : |:   | ||        | :: |   ::||
Db       2048 STNYTVATGNFHPSGSPLGPSSGSTGESYGLSPLRPPSVLPPPAPDGSLPY---LSHGAS 2104

Qy       1235 ADVLYKGTITRIIGEDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKED- 1293
                    | : :   : |     |   ||    :   | :   :|   |:| |:  |:
Db       2105 Q---RSGITSPVEKREDPG---TGMGSSLATAEL--PGTQDPGMS---GLSQTELEKQRQ 2153

Qy       1294 ------------GRSSSGPPHETAAPKRTYDMMEGRVGRAISSASIE------------ 1328
                         |::   ||||          |  |    || : |
Db       2154 RQRLRELLIRQQIQRNTLRQEKETAAAAAGAVGPPGSWGAEPSSPAFEQLSRGQTPFAGT 2213

Qy       1329 ----GLMGRAIPPERHSPHHLKEQHHIRGSITQGIPRSYVEAQEDYLRREAKLLKREGTP 1384
                   |:|  :|| :          : | |    | |:    :| | |             |
Db       2214 QDKSSLVG--LPPSK-----------LSGPILG--PGSF--PSDDRLSR----------P 2246

Qy       1385 PPP--PPSRDL--------TEAYKTQA--LGPLKLKPAHEGL-------VATVKEAGRSI 1425
                |||  | | |:        ::|:  :|   | | |:   : |         ||       |
```





RESULT 21
US-09-825-751A-79
; Sequence 79, Application US/09825751A
; Publication No. US20030065140A1
; GENERAL INFORMATION:
; APPLICANT: CuraGen Corporation
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Fernandes, Elma R
; APPLICANT: Taupier, Raymond J
; APPLICANT: Quinn, Kerry E
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Rastelli, Luca
; APPLICANT: Herrman, John L
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-750
; CURRENT APPLICATION NUMBER: US/09/825,751A
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/194,314
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: 60/225,693
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 79
; LENGTH: 1151
; TYPE: PRT
; ORGANISM: Gallus gallus
US-09-825-751A-79

Query Match 3.3%; Score 436; DB 12; Length 1151;
Best Local Similarity 19.7%; Pred. No. 8.6e-12;
Matches 346; Conservative 154; Mismatches 548; Indels 706; Gaps 77;



```
Db        399 -----GAVPKAST------GTTPAAAPQQPV----------PKAAPVTPPSPQQAVPR-- 435
Qy       1178 AGPPESLGVPTAQEASVLRGTALGSVPGGSITKGIPSTR-VPSDSAITYRGSITHGTPAD 1236
Db        436 --------AATAAAAPV--------TPQQPVTKAATTTNATPPPQPIPKAATTTTATPV- 478
Qy       1237 VLYKGTITRIIGEDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRS 1296
Db        479 --------------TP-------QQPIPKA--------------------------GTD 490
Qy       1297 SSGPPHETAAPKRTYDMMEGRVGRAISSASIEGLMGRAIPPERHSPHHLKEQHHIRGSIT 1356
Db        491 AAPPP---AVPKAPSD-------------------GRAATP------------------- 509
Qy       1357 QGIPRSYVEAQEDYLRREAKLLKREGTPPPPPPSRDLTEAYKTQALGPLKLKPAHEGLVA 1416
Db        510 -GVPNAATDPQK--------------PPPTPQS------------------------- 527
Qy       1417 TVKEAGRSIHEIPREELRHTPELPLAPRPLKEGSITQGTP---LKYDTGASTTGSKKHDV 1473
Db        528 -----------VPSAVTEPKPQPRAAPPPSNEATPAVPSPSPNLK-------------- 561
Qy       1474 RSLIGSPGRTFPPVHPLDVMADARALERACYEESLKSRPGTASSSGGSIARGAPVIVPEL 1533
Db        562 -----SPLPTIP--KPVPLMA------------LTPQPVTAQ-----------MVTQL 589
Qy       1534 GKPRQSPLTYEDHGAPFAGHLPRGSPVTMREPTPRLQEGSLSSSKASQDRKLTSTPREIA 1593
Db        590 AATKPSPI------------VPKASPKALMTPPPP-PPGLPRALAAAKLLGLPSSP--VA 634
Qy       1594 KSPHSTVPEHHPHPISPYEHLLRGVSGVDLYRSHIPLAFDPTSIPRGIPLDAAAAYYLPR 1653
Db        635 SAMHAKVTP-RPLPASP----------------VPMAASPASLGP----DAARV----- 667
Qy       1654 HLAPNPTYPHLYPPYLIRGYPDTAALENRQTIINDYITSQQMHHNTATAMAQRADMLRGL 1713
Db        668 ALATNAASPGAKP-------------------------EAAGGNGTLMAPMG------ 694
Qy       1714 SPRESSLALNYAAGPRGIIDLSQVPHLPVL---VPPTPGTPATAMDRLAYLPTAPQPFSS 1770
Db        695 -------AANTQMAPIGAAGAAQTAPMGAAHTHVSPMGAGGATQM--------SPTGAAN 739
Qy       1771 RHSSSPLSPGGPTHLTKPTTTSSSERERDRDRERDRDREREKSILTSTTTVEHAPIWRPG 1830
Db        740 TH-MSPIGAGGATQMS-PMGAANTQMSP----------------MGATTTQMSPMGAAA 780
Qy       1831 TEQSSGSSGSSGGGGGSSSRPASHSHAHQHSPISPRTQDALQQRPSVLHNTGMKGIITAV 1890
Db        781 TTQPSPM-----GAAATQVTATSAGNTMQVSPMGAATP---PQTPSV-------GAATTP 825
Qy       1891 EPSKPTVLRSTSTSSPVRPAATFPPATHCPLGGTLDGVYPTLMEPVLLPKEAPRVARPER 1950
Db        826 QPS--PMGAATTLMSPMGAATTPQPS---PMGAV------TTQPP---PMAATNTTQPPP 871
Qy       1951 PRADT------GHAFLAKPPARSGLEPASSPSKGSEPRPLVPPVSGHATIARTPAKNLAP 2004
Db        872 MAASTPQSTPMGAATTTQSPPMGATTTQSPPMGASTPQ--APP-----TVAGSP------ 918
Qy       2005 HHASPDPPAPPASASDPHREKTQSKPFSIQELELRSLGYHGSSYSPEGVEPVSPVSSPSL 2064
Db        919 ---TPPPPIPPSPTA-----QTSPQPMS----------------------KSPPPDP-- 945
Qy       2065 THDKGLPKHLEELDKSHLEGELRPKQPGPVKLGGEAAHLPHLRPLPESQPSSSPLLQTAP 2124
Db        946 ----------------------PKAPSAAAQTSPAAHVAN---------------ASP 966
Qy       2125 GVKGHQRVVTLAQHISEVITQDYTRHHPQQLSAPLPAPLYSFPGASCPVLDLRRPPSDLY 2184
Db        967 GV---------------------------TAVSPAPI-------------------- 976
Qy       2185 LPPPDHGAPARGSPHSEGGKRSPEPNKTSVLGGGEDGIEPVSPPEG---MTEPGHSRSAV 2241
Db        977 -------GVTEASPSADGARLSPGPTAAT------DG--PKASPAATADVTEAATDVTAA 1021

Qy       2242 YPLLYRDGEQTEPSRMGSKSPGNTSQPPAFFSKLTESNSAMVKSKKQEINKKLNTHNRNE 2301
Db       1022 ATAVPAEAAPTKAKRSSSSSSSSS------SSSSSSSSSSSSSDSDSSSSSSESNPAS 1075

Qy       2302 PEYNISQPGTEIFNMPAITGTGLMTYRSQAVQEHASTNMGLEAIIRKALMGKYDQWEESP 2361
Db       1076 P-------------APAVG---------------------------------DGQQQMT 1088

Qy       2362 PLSANAFNPLNASASLPAAMPITAADGRSDHTLTSPGGGGKAKVSGRPSSRKAKSPAPGL 2421
Db       1089 PGAAQSVPPVTEAAVQEAAAAAAAAAG-------------AEREGRPTRRKKRTRSSSS 1134

Qy       2422 ASGDRPPSVSSVHS 2435
Db       1135 SSSSSSSSSSSSS 1148
```

RESULT 22
US-09-738-973-425
; Sequence 425, Application US/09738973
; Patent No. US20020110563A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Fling, Steven P.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Algate, Paul A.
; APPLICANT: Secrist, Heather
; APPLICANT: Indirias, Carol Yoseph
; APPLICANT: Benson, Darin R.
; APPLICANT: Elliot, Mark
; APPLICANT: Mannion, Jane
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C9
; CURRENT APPLICATION NUMBER: US/09/738,973
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 587
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 425
; LENGTH: 4019
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-738-973-425

Query Match 3.3%; Score 436; DB 9; Length 4019;
Best Local Similarity 18.3%; Pred. No. 3.8e-11;
Matches 555; Conservative 314; Mismatches 1010; Indels 1152; Gaps 142;

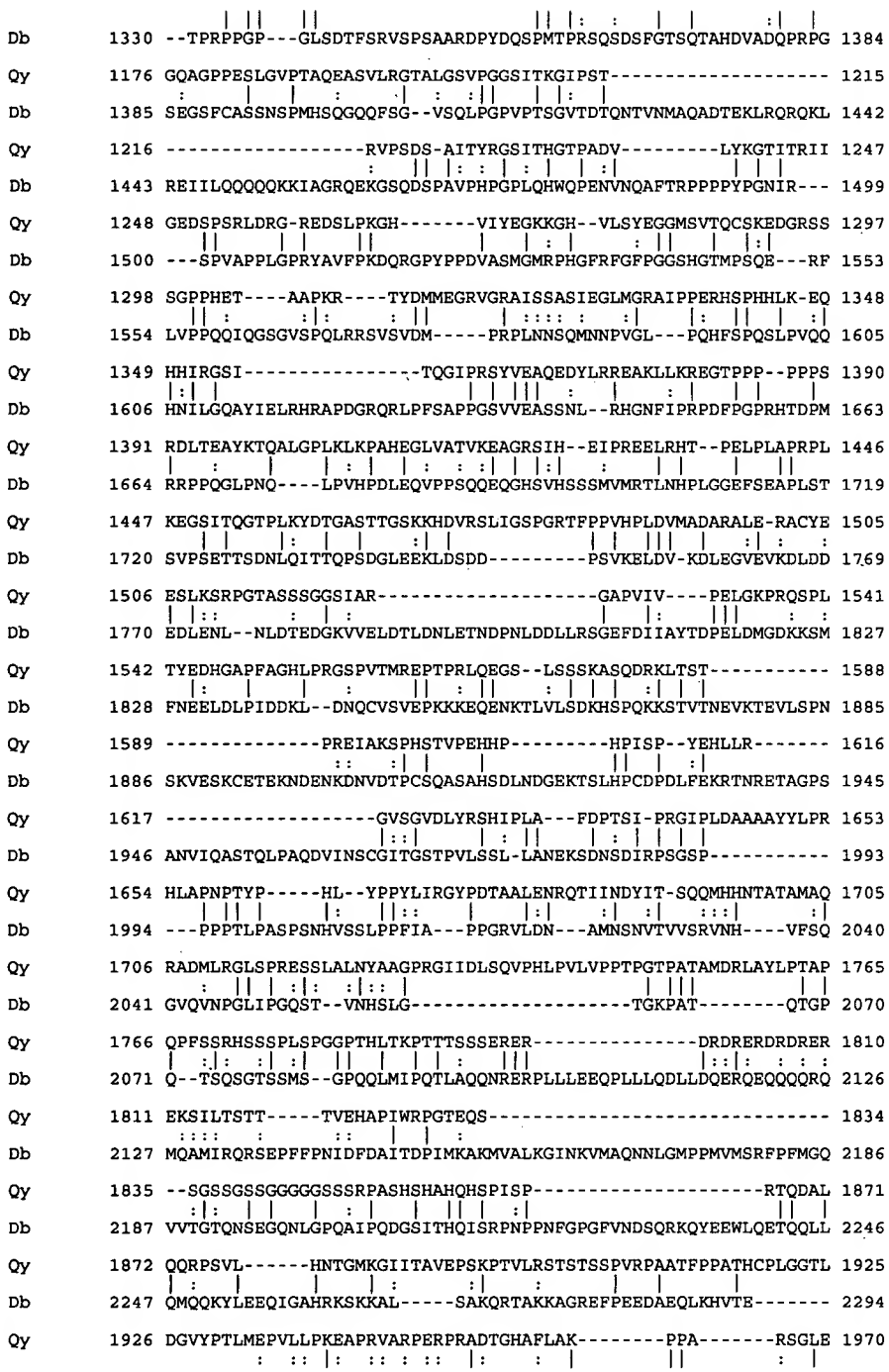
```
Qy         34 HTDVGLLEYQHHSRDYASHLSPGSIIQP---QRRRPSLLSEFQPGNERSQELHLRPESHS 90
Db        513 HSDIGPVTD-----------DPSSLPQPNVNQSSRP--LSEEQLDGILSPELDKMVTDGA 559
Qy         91 YL------PELGKSEMEFIESKRPRLELLPDPLLRPSPLLATGQPAGSEDLTKDRSLTGK 144
Db        560 ILGKLYKIPELGGKDVEDL-------------------FTAVLSPANT------------ 588
Qy        145 LEPVSPPSPPHTDPELELVPPRLSKEELIQNMDRVDREITMVEQQISKLKKKQQQLEEEA 204
Db        589 -QPTPLPQPP---PPTQLLP--------IHNQDAFSR-MPLMNGLI-------------G 622
Qy        205 AKPPEPEKPVSPPPIESKHRSLVQIIYDENRKKAEAAHRILEGLGPQVELPLYNQPSDTR 264
Db        623 SSPHLPHNSLPPGSGLGTFSAIAQSSYPDARDKNSAFN------------PMASDPNNS- 669
Qy        265 QYHENIKINQAMRKKLILYFKRRNHARKQWKQKFCQRYDQLMEALEKKVERIENNPRRRA 324
Db        670 ----------------------------WTSS------------APTVEG-ENDTMSNA 687
Qy        325 KESKVREYYEKQFPEIRKQRELQERMQSRVGQRGSGLSMSAARSEHEVSEIIDGLSEQEN 384
```

```
Qy       2062 PSLTHDKGLPKHLEEL-----DKSHLEGELRP-----------KQPGPVKLGGEAAHLP 2104
Db       2527 S--------PQHQQQMGQVLQQQNIQQGSINSPSTQTFMQTNERRQVGPPSFVPDSPSIP 2578

Qy       2105 HLRPLPES-----------------QPSSSPLLQTAPGVKGHQRVVTLAQHISEVITQD 2146
Db       2579 VGSPNFSSVKQGHGNLSGTSFQQSPVRPSFTPALPAAPPV----------ANSSLPCGQD 2628

Qy       2147 YTRHHPQQLSAPLPAPLYSFPGASCPVLDLRRPPSDLYLPPPDHGAPARGSPHSEGGKRS 2206
Db       2629 STITHG----------HSYPGSTQSLIQLY---SDII------------PEEKGKKKR 2661

Qy       2207 PEPNKTSVLGGGEDGIEPVSPPEGMT---EPGHSRSAVYPLLYRDGEQTEPSRMGSKSPG 2263
Db       2662 TRKKKRD--DDAESTKAPSTPHSDITAPPTPGISETTSTPAVSTPSELPQQADQESVEPV 2719

Qy       2264 NTSQPPAFFSKL-TESNSAMVKSKKQEINKKLNTHNRNEPEYNISQPGTEIFNMPAITGT 2322
Db       2720 GPSTPNMAAGQLCTELENKLP-------NSDFSQATPNQQTYANSEVDKLSMETPAKT-- 2770

Qy       2323 GLMTYRSQAVQEHASTNMGLEAIIRKALMGKYDQWEESPPLSANAFNPLNASASLPAAMP 2382
Db       2771 --------------EEIKLEKAETESCPG-----QEEPKLEEQNGSKVEGNA---VACP 2807

Qy       2383 ITAADGRSDHTLTSP---GGGGKAKVSGRPSSRKAKSPAPGLASGDRPPSVSSVHSEGDC 2439
Db       2808 VSSAQS-PPHSAGAPAAKGDSGNELLKHLLKNKKSSS-----LLNQKPE--GSICSEDDC 2859

Qy       2440 NRRTPLTNRVWEDRPSSAGSTPFPYNPLIMRLQAGVMASPPPPGLPAGS---------- 2488
Db       2860 TKDNKL---VEKQNPAEGLQT------LGAQMQGGFGCGNQLPKTDGGSETKKQRSKRTQ 2910

Qy       2489 --GPLAGPH---HAWDEEPKPLLCSQYETLS 2514
Db       2911 RTGEKAAPRSKKRKKDEEEKQAMYSSTDTFT 2941
```

RESULT 23
US-09-854-133-425
; Sequence 425, Application US/09854133
; Publication No. US20020183499A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Henderson, Robert A.
; APPLICANT: Benson, Darin R.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C10
; CURRENT APPLICATION NUMBER: US/09/854,133
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 425
; LENGTH: 4019
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-854-133-425

Query Match 3.3%; Score 436; DB 9; Length 4019;
Best Local Similarity 18.3%; Pred. No. 3.8e-11;
Matches 555; Conservative 314; Mismatches 1010; Indels 1152; Gaps 142;

```
Qy         34 HTDVGLLEYQHHSRDYASHLSPGSIIQP---QRRRPSLLSEFQPGNERSQELHLRPESHS 90
Db        513 HSDIGPVTD-----------DPSSLPQPNVNQSSRP--LSEEQLDGILSPELDKMVTDGA 559

Qy         91 YL------PELGKSEMEFIESKRPRLELLPDPLLRPSPLLATGQPAGSEDLTKDRSLTGK 144
Db        560 ILGKLYKIPELGGKDVEDL------------------FTAVLSPANT------------ 588

Qy        145 LEPVSPPSPPHTDPELELVPPRLSKEELIQNMDRVDREITMVEQQISKLKKKQQQLEEEA 204
Db        589 -QPTPLPQPP---PPTQLLP--------IHNQDAFSR-MPLMNGLI------------G 622

Qy        205 AKPPEPEKPVSPPPIESKHRSLVQIIYDENRKKAEAAHRILEGLGPQVELPLYNQPSDTR 264
Db        623 SSPHLPHNSLPPGSGLGTFSAIAQSSYPDARDKNSAFN-----------PMASDPNNS- 669

Qy        265 QYHENIKINQAMRKKLILYFKRRNHARKQWKQKFCQRYDQLMEALEKKVERIENNPRRRA 324
Db        670 ----------------------------WTSS-----------APTVEG-ENDTMSNA 687

Qy        325 KESKVREYYEKQFPEIRKQRELQERMQSRVGQRGSGLSMSAARSEHEVSEIIDGLSEQEN 384
Db        688 QRSTLK--WEK---------------------------------------------- 696

Qy        385 LEKQMRQLAVIPPMLYDADQQRIKFINMNGLMADPMKVYKDRQVMNMWSEQEKETFREKF 444
Db        697 -EEALGEMATVAPVLY----TNINFPNLKEEFPD--WTTRVKQIAKLWRK---------- 739

Qy        445 MQHPKNFGLIASFLERKTVAECVLYYYLTKKNENYKSLVRRSYRRRGKSQQQQQQQQQQQ 504
Db        740 ----------ASSQERAP---------YVQKARDN-------RAALRINKVQMSNDSMKRQQ 775

Qy        505 QQQQQQPMPRSSQEEKDEKEKEKEAEKEEEKPEVENDKEDLLKEKTDDTSGEDNDEKEAV 564
Db        776 QQDSIDPSSRIDSELFKDPLKQRESEHEQE---------------------------- 805

Qy        565 ASKGRKTANSQGRRKGRI--TRSMANEANSEEAITPQQSAELASMELNESSRWTEEEMET 622
Db        806 -WKFRQQMRQKSKQQAKIEATQKLEQVKNEQQ---QQQQQQFGSQHL-----LVQSGSDT 856

Qy        623 AKKGLLEHGRNWSAIARMVGSKTVSQCKNFYFNYKKRQNLDEILQQHKLKMEKERNARRK 682
Db        857 PSSGI------QSPLTPQPGNGNMSPAQSF----------------HKELFTKQ------ 888

Qy        683 KKKAPAAASEEAAF--------PPV-----VEDEEMEASGVSGNEEEMVEEAEALHASGN 729
Db        889 PPSTPTSTSSDDVFVKPQAPPPPAPSRIPIQDSLSQA---------------------- 926

Qy        730 EVPRGECSGPATVNNSSDTESIPSPHTEAAKDTGQNGPKPPATLGADGPPPGPPTPP--- 786
Db        927 ----------------QTSQPPSPQVFS---PGSSNSRPPSPMDPYAKMVGTPRPPPVG 966

Qy        787 ----RRTSRAPIEP-TPASEA--------TGAPTPPPAPP---------SPSAPPPVVPK 824
Db        967 HSFSRRNSAAPVENCTPLSSVSRPLQMNETTANRPSPVRDLCSSSTTNNDPYAKPPDTP- 1025

Qy        825 EEKEEETAAAPPVEEGEEQKPPAAEELAVDTGKAEEPVKSECTEEAEEGPAKGKDAEAAE 884
Db       1026 ------------------RPVMTDQFPKSLGLSRSPVVSEQT-------AKG---PIAA 1056

Qy        885 ATAEGALKAEKKEGGSGRATTAKSSGAPQDSDSSATCSADEVDEAEGGDKNRLLSPRPSL 944
Db       1057 GTSDHFTKP------SPRADVFQRQRIP-DSYARPLLTPAPLDSGPGPFKTP-MQPPPSS 1108

Qy        945 LTPTGDPRANASPQKPLDLKQLKQRAAAIPPIQVTKVHEPPREDAAPTKPAPPAPPPPQN 1004
Db       1109 QDPYGSV-SQASRRLSVD----PYERPALTPRPIDNFSHNQSNDPYSQPPLTPHPAVNES 1163

Qy       1005 LQPESDAPQQPGSSPRGKSR--------SPAPPADKEAFAAEAQKLPGDPPCWTSGLPFP 1056
Db       1164 FAHPSRAFSQPGTISRPTSQDPYSQPPGTPRPVVDSYSQSSGTARSNTDPYSQPPGTPRP 1223

Qy       1057 --VPPREVIKASPHAPDPSAFS---------------YA-PPGHPLP-LGLHDTARPVLP 1097
Db       1224 TTVDP---YSQQPQTPRPSTQTDLFVTPVTNQRHSDPYAHPPGTPRPGISVPYSQPPATP 1280

Qy       1098 RP-------------PTI-SNPPPLISSAKH-----PSVLERQIGAISQGMSVQLHVPYS 1138
Db       1281 RPRISEGFTRSSMTRPVLMPNQDPFLQAAQNRGPALPGPLVRPPDTCSQ----------- 1329

Qy       1139 EHAKAPVGPVTMGL----------------PLPMDPKKLAPFSGVKQ------EQLSPR 1175
```

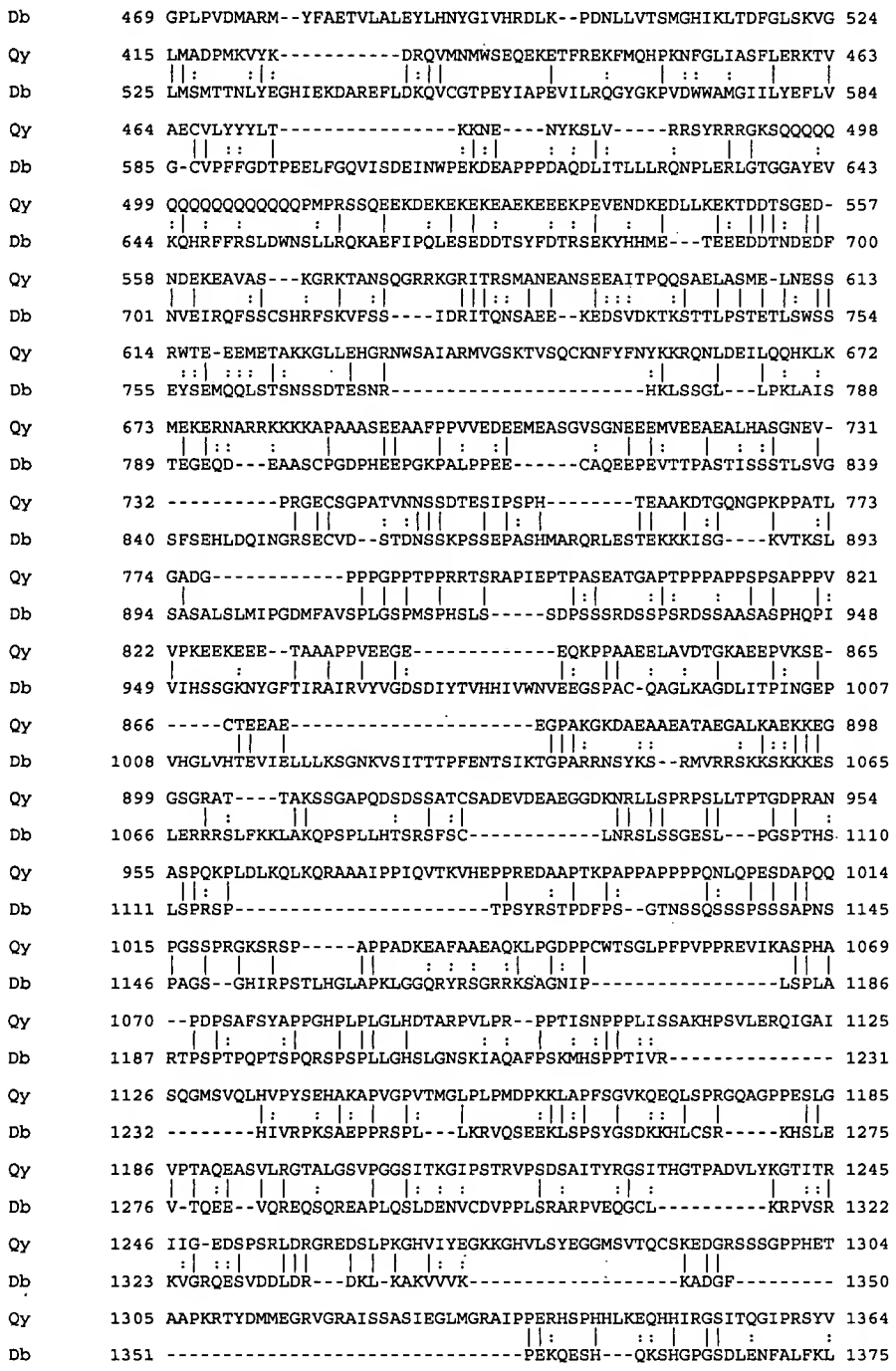


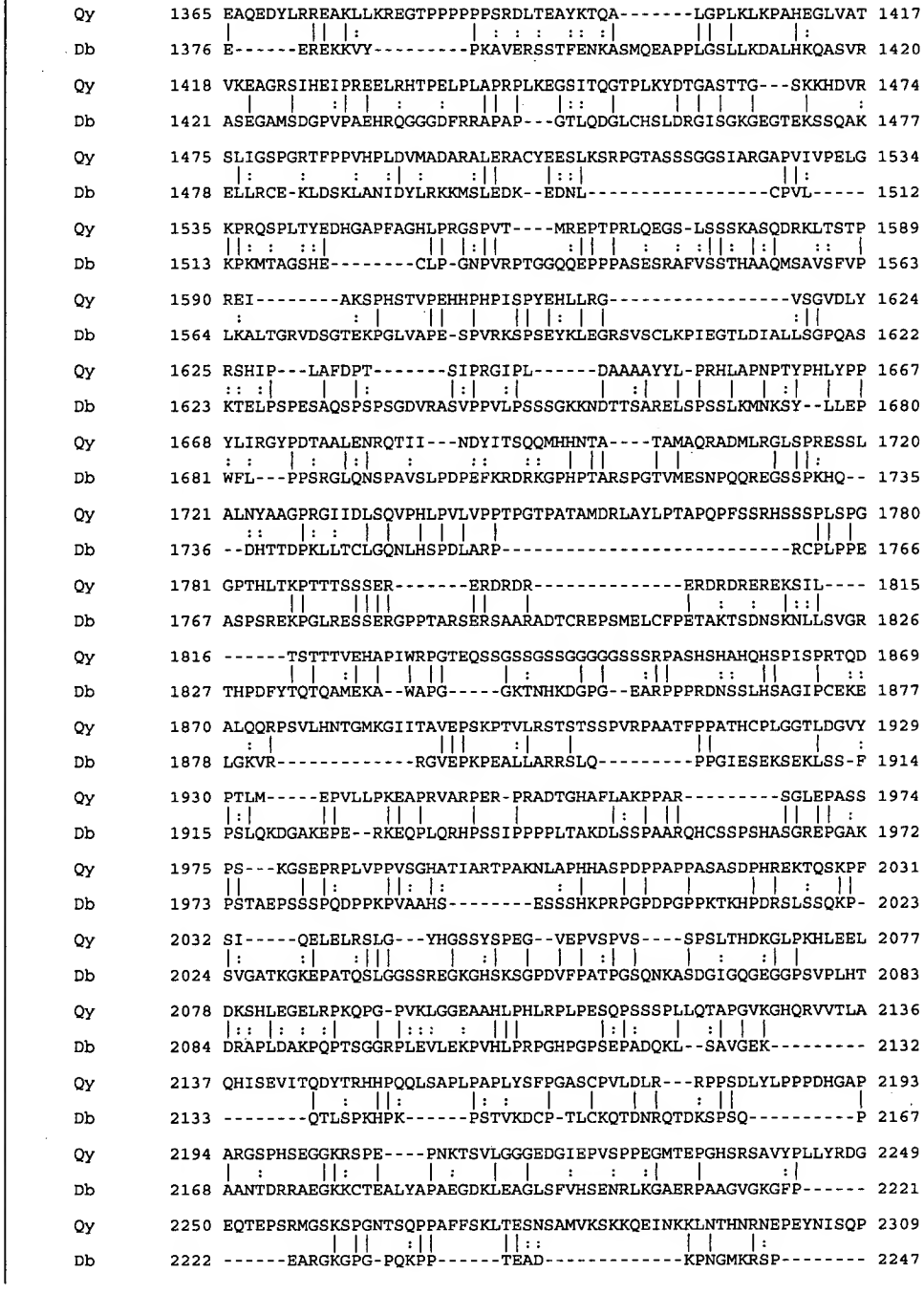
RESULT 24
US-10-144-649A-425
; Sequence 425, Application US/10144649A
; Publication No. US20030118599A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Algate, Paul A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C11
; CURRENT APPLICATION NUMBER: US/10/144,649A
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 749
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 425
; LENGTH: 4019
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-649A-425

Query Match 3.3%; Score 436; DB 14; Length 4019;
Best Local Similarity 18.3%; Pred. No. 3.8e-11;
Matches 555; Conservative 314; Mismatches 1010; Indels 1152; Gaps 142;

```
Qy        34 HTDVGLLEYQHHSRDYASHLSPGSIIQP---QRRRPSLLSEFQPGNERSQELHLRPESHS  90
Db       513 HSDIGPVTD-----------DPSSLPQPNVNQSSRP--LSEEQLDGILSPELDKMVTDGA  559

Qy        91 YL------PELGKSEMEFIESKRPRLELLPDPLLRPSPLLATGQPAGSEDLTKDRSLTGK  144
Db       560 ILGKLYKIPELGGKDVEDL------------------FTAVLSPANT------------  588

Qy       145 LEPVSPPSPPHTDPELELVPPRLSKEELIQNMDRVDREITMVEQQISKLKKKQQQLEEEA  204
Db       589 -QPTPLPQPP---PPTQLLP--------IHNQDAFSR-MPLMNGLI-------------G  622

Qy       205 AKPPEPEKPVSPPPIESKHRSLVQIIYDENRKKAEAAHRILEGLGPQVELPLYNQPSDTR  264
Db       623 SSPHLPHNSLPPGSGLGTFSAIAQSSYPDARDKNSAFN-----------PMASDPNNS-  669

Qy       265 QYHENIKINQAMRKKLILYFKRRNHARKQWKQKFCQRYDQLMEALEKKVERIENNPRRRA  324
Db       670 ----------------------------WTSS------------APTVEG-ENDTMSNA  687

Qy       325 KESKVREYYEKQFPEIRKQRELQERMQSRVGQRGSGLSMSAARSEHEVSEIIDGLSEQEN  384
Db       688 QRSTLK--WEK-----------------------------------------------------  696

Qy       385 LEKQMRQLAVIPPMLYDADQQRIKFINMNGLMADPMKVYKDRQVMNMWSEQEKETFREKF  444
Db       697 -EEALGEMATVAPVLY----TNINFPNLKEEFPD--WTTRVKQIAKLWRK----------  739

Qy       445 MQHPKNFGLIASFLERKTVAECVLYYYLTKKNENYKSLVRRSYRRRGKSQQQQQQQQQQQ  504
Db       740 ----------ASSQERAP--------YVQKARDN------RAALRINKVQMSNDSMKRQQ  775

Qy       505 QQQQQQPMPRSSQEEKDEKEKEKEAEKEEEKPEVENDKEDLLKEKTDDTSGEDNDEKEAV  564
Db       776 QQDSIDPSSRIDSELFKDPLKQRESEHEQE-----------------------------  805

Qy       565 ASKGRKTANSQGRRKGRI--TRSMANEANSEEAITPQQSAELASMELNESSRWTEEEMET  622
Db       806 -WKFRQQMRQKSKQQAKIEATQKLEQVKNEQQ---QQQQQQFGSQHL-----LVQSGSDT  856

Qy       623 AKKGLLEHGRNWSAIARMVGSKTVSQCKNFYFNYKKRQNLDEILQQHKLKMEKERNARRK  682
Db       857 PSSGI------QSPLTPQPGNGNMSPAQSF---------------HKELFTKQ------  888

Qy       683 KKKAPAAASEEAAF--------PPV-----VEDEEMEASGVSGNEEEMVEEAEALHASGN  729
Db       889 PPSTPTSTSSDDVFVKPQAPPPPPAPSRIPIQDSLSQA----------------------  926

Qy       730 EVPRGECSGPATVNNSSDTESIPSPHTEAAKDTGQNGPKPPATLGADGPPPGPPTPP---  786
Db       927 ----------------QTSQPPSPQVFS---PGSSNSRPPSPMDPYAKMVGTPRPPPVG  966

Qy       787 ----RRTSRAPIEP-TPASEA--------TGAPTPPPAPP---------SPSAPPPVVPK  824
Db       967 HSFSRRNSAAPVENCTPLSSVSRPLQMNETTANRPSPVRDLCSSSTTNNDPYAKPPDTP-  1025

Qy       825 EEKEEETAAAPPVEEGEEQKPPAAEELAVDTGKAEEPVKSECTEEAEEGPAKGKDAEAAE  884
Db      1026 ------------------RPVMTDQFPKSLGLSRSPVVSEQT-------AKG---PIAA  1056

Qy       885 ATAEGALKAEKKEGGSGRATTAKSSGAPQDSDSSATCSADEVDEAEGGDKNRLLSPRPSL  944
Db      1057 GTSDHFTKP------SPRADVFQRQRIP-DSYARPLLTPAPLDSGPGPFKTP-MQPPPSS  1108

Qy       945 LTPTGDPRANASPQKPLDLKQLKQRAAAIPPIQVTKVHEPPREDAAPTKPAPPAPPPPQN  1004
Db      1109 QDPYGSV-SQASRRLSVD----PYERPALTPRPIDNFSHNQSNDPYSQPPLTPHPAVNES  1163

Qy      1005 LQPESDAPQQPGSSPRGKSR--------SPAPPADKEAFAAEAQKLPGDPPCWTSGLPFP  1056
Db      1164 FAHPSRAFSQPGTISRPTSQDPYSQPPGTPRPVVDSYSQSSGTARSNTDPYSQPPGTPRP  1223

Qy      1057 --VPPREVIKASPHAPDPSAFS--------------YA-PPGHPLP-LGLHDTARPVLP  1097
Db      1224 TTVDP---YSQQPQTPRPSTQTDLFVTPVTNQRHSDPYAHPPGTPRPGISVPYSQPPATP  1280

Qy      1098 RP-------------PTI-SNPPPLISSAKH-----PSVLERQIGAISQGMSVQLHVPYS  1138
Db      1281 RPRISEGFTRSSMTRPVLMPNQDPFLQAAQNRGPALPGPLVRPPDTCSQ-----------  1329

Qy      1139 EHAKAPVGPVTMGL-----------------PLPMDPKKLAPFSGVKQ------EQLSPR  1175
Db      1330 --TPRPPGP---GLSDTFSRVSPSAARDPYDQSPMTPRSQSDSFGTSQTAHDVADQPRPG  1384

Qy      1176 GQAGPPESLGVPTAQEASVLRGTALGSVPGGSITKGIPST--------------------  1215
Db      1385 SEGSFCASSNSPMHSQGQQFSG--VSQLPGPVPTSGVTDTQNTVNMAQADTEKLRQRQKL  1442

Qy      1216 ------------------RVPSDS-AITYRGSITHGTPADV---------LYKGTITRII  1247
Db      1443 REIILQQQQQKKIAGRQEKGSQDSPAVPHPGPLQHWQPENVNQAFTRPPPPYPGNIR---  1499

Qy      1248 GEDSPSRLDRG-REDSLPKGH-------VIYEGKKGH--VLSYEGGMSVTQCSKEDGRSS  1297
Db      1500 ---SPVAPPLGPRYAVFPKDQRGPYPPDVASMGMRPHGFRFGFPGGSHGTMPSQE---RF  1553

Qy      1298 SGPPHET----AAPKR----TYDMMEGRVGRAISSASIEGLMGRAIPPERHSPHHLK-EQ  1348
Db      1554 LVPPQQIQGSGVSPQLRRSVSVDM-----PRPLNNSQMNNPVGL---PQHFSPQSLPVQQ  1605

Qy      1349 HHIRGSI----------------TQGIPRSYVEAQEDYLRREAKLLKREGTPPP--PPPS  1390
Db      1606 HNILGQAYIELRHRAPDGRQRLPFSAPPGSVVEASSNL--RHGNFIPRPDFPGPRHTDPM  1663

Qy      1391 RDLTEAYKTQALGPLKLKPAHEGLVATVKEAGRSIH--EIPREELRHT--PELPLAPRPL  1446
Db      1664 RRPPQGLPNQ----LPVHPDLEQVPPSQQEQGHSVHSSSMVMRTLNHPLGGEFSEAPLST  1719

Qy      1447 KEGSITQGTPLKYDTGASTTGSKKHDVRSLIGSPGRTFPPVHPLDVMADARALE-RACYE  1505
Db      1720 SVPSETTSDNLQITTQPSDGLEEKLDSDD---------PSVKELDV-KDLEGVEVKDLDD  1769

Qy      1506 ESLKSRPGTASSSGGSIAR--------------------GAPVIV----PELGKPRQSPL  1541
Db      1770 EDLENL--NLDTEDGKVVELDTLDNLETNDPNLDDLLRSGEFDIIAYTDPELDMGDKKSM  1827

Qy      1542 TYEDHGAPFAGHLPRGSPVTMREPTPRLQEGS--LSSSKASQDRKLTST-----------  1588
Db      1828 FNEELDLPIDDKL--DNQCVSVEPKKKEQENKTLVLSDKHSPQKKSTVTNEVKTEVLSPN  1885

Qy      1589 --------------PREIAKSPHSTVPEHHP---------HPISP--YEHLLR-------  1616
Db      1886 SKVESKCETEKNDENKDNVDTPCSQASAHSDLNDGEKTSLHPCDPDLFEKRTNRETAGPS  1945

Qy      1617 -------------------GVSGVDLYRSHIPLA---FDPTSI-PRGIPLDAAAAYYLPR  1653
Db      1946 ANVIQASTQLPAQDVINSCGITGSTPVLSSL-LANEKSDNSDIRPSGSP-----------  1993

Qy      1654 HLAPNPTYP-----HL--YPPYLIRGYPDTAALENRQTIINDYIT-SQQMHHNTATAMAQ  1705
Db      1994 ---PPPTLPASPSNHVSSLPPFIA---PPGRVLDN---AMNSNVTVVSRVNH----VFSQ  2040

Qy      1706 RADMLRGLSPRESSLALNYAAGPRGIIDLSQVPHLPVLVPPTPGTPATAMDRLAYLPTAP  1765
Db      2041 GVQVNPGLIPGQST--VNHSLG--------------------TGKPAT--------QTGP  2070

Qy      1766 QPFSSRHSSSPLSPGGPTHLTKPTTTSSSERER---------------DRDRERDRDRER  1810
Db      2071 Q--TSQSGTSSMS--GPQQLMIPQTLAQQNRERPLLLEEQPLLLQDLLDQERQEQQQQRQ  2126

Qy      1811 EKSILTSTT-----TVEHAPIWRPGTEQS-----------------------------  1834
```

```
Db       2127 MQAMIRQRSEPFFPNIDFDAITDPIMKAKMVALKGINKVMAQNNLGMPPMVMSRFPFMGQ 2186

Qy       1835 --SGSSGSSGGGGGSSSRPASHSHAHQHSPISP--------------------RTQDAL 1871

Db       2187 VVTGTQNSEGQNLGPQAIPQDGSITHQISRPNPPNFGPGFVNDSQRKQYEEWLQETQQLL 2246

Qy       1872 QQRPSVL------HNTGMKGIITAVEPSKPTVLRSTSTSSPVRPAATFPPATHCPLGGTL 1925

Db       2247 QMQQKYLEEQIGAHRKSKKAL-----SAKQRTAKKAGREFPEEDAEQLKHVTE------- 2294

Qy       1926 DGVYPTLMEPVLLPKEAPRVARPERPRADTGHAFLAK--------PPA-------RSGLE 1970

Db       2295 --------QQSMVQKQLEQIRKQQKEHAELIEDYRIKQQQQCAMAPPTMMPSVQPQPPLI 2346

Qy       1971 PASSPSKGSEPR-PLVPP----------VSGHATIARTPA-----KNLAPHHASPDPP-- 2012

Db       2347 PGATPPTMSQPTFPMVPQQLQHQQHTTVISGHTSPVRMPSLPGWQPNSAPAHLPLNPPRI 2406

Qy       2013 --------------------APPASASDPHREKTQSKPFS------------------- 2032

Db       2407 QPPIAQLPIKTCTPAPGTVSNANPQSGPPPRVEFDDNNPFSESFQERERKERLREQQERQ 2466

Qy       2033 --------------IQELELRSLGYHGSSYSP-----------------EGVEPVSPVSS 2061

Db       2467 RIQLMQEVDRQRALQQRMEMEQHGMVGSEISSSRTSVSQIPFYSSDLPCDFMQPLGPLQQ 2526

Qy       2062 PSLTHDKGLPKHLEEL-----DKSHLEGELRP------------KQPGPVKLGGEAAHLP 2104

Db       2527 S--------PQHQQQMGQVLQQQNIQQGSINSPSTQTFMQTNERRQVGPPSFVPDSPSIP 2578

Qy       2105 HLRPLPES------------------QPSSSPLLQTAPGVKGHQRVVTLAQHISEVITQD 2146

Db       2579 VGSPNFSSVKQGHGNLSGTSFQQSPVRPSFTPALPAAPPV----------ANSSLPCGQD 2628

Qy       2147 YTRHHPQQLSAPLPAPLYSFPGASCPVLDLRRPPSDLYLPPPDHGAPARGSPHSEGGKRS 2206

Db       2629 STITHG-----------HSYPGSTQSLIQLY---SDII-------------PEEKGKKKR 2661

Qy       2207 PEPNKTSVLGGGEDGIEPVSPPEGMT---EPGHSRSAVYPLLYRDGEQTEPSRMGSKSPG 2263

Db       2662 TRKKKRD--DDAESTKAPSTPHSDITAPPTPGISETTSTPAVSTPSELPQQADQESVEPV 2719

Qy       2264 NTSQPPAFFSKL-TESNSAMVKSKKQEINKKLNTHNRNEPEYNISQPGTEIFNMPAITGT 2322

Db       2720 GPSTPNMAAGQLCTELENKLP-------NSDFSQATPNQQTYANSEVDKLSMETPAKT-- 2770

Qy       2323 GLMTYRSQAVQEHASTNMGLEAIIRKALMGKYDQWEESPPLSANAFNPLNASASLPAAMP 2382

Db       2771 --------------EEIKLEKAETESCPG-----QEEPKLEEQNGSKVEGNA---VACP 2807

Qy       2383 ITAADGRSDHTLTSP---GGGGKAKVSGRPSSRKAKSPAPGLASGDRPPSVSSVHSEGDC 2439

Db       2808 VSSAQS-PPHSAGAPAAKGDSGNELLKHLLKNKKSSS-----LLNQKPE--GSICSEDDC 2859

Qy       2440 NRRTPLTNRVWEDRPSSAGSTPFPYNPLIMRLQAGVMASPPPPGLPAGS----------- 2488

Db       2860 TKDNKL---VEKQNPAEGLQT------LGAQMQGGFGCGNQLPKTDGGSETKKQRSKRTQ 2910

Qy       2489 --GPLAGPH---HAWDEEPKPLLCSQYETLS 2514

Db       2911 RTGEKAAPRSKKRKKDEEEKQAMYSSTDTFT 2941


RESULT 25
US-10-362-892-8
; Sequence 8, Application US/10362892
; Publication No. US20040038881A1
; GENERAL INFORMATION:
;  APPLICANT: INCYTE GENOMICS, INC.; BANDMAN, Olga
;  APPLICANT:  NGUYEN, Danniel B.; WALIA, Narinder K.
;  APPLICANT:  HAFALIA, April J.A.; YAO, Monique G.
;  APPLICANT:  GANDHI, Ameena R.; GURURAJAN, Rajagopal
;  APPLICANT:  DING, Li; PATTERSON, Chandra S.
;  APPLICANT:  YUE, Henry; BAUGHN, Mariah R.
;  APPLICANT:  TRIBOULEY, Catherine M.; THORNTON, Michael B.
;  APPLICANT:  ELLIOTT, Vicki S.; LU, Yan
;  APPLICANT:  ISON, Craig H.; AU-YOUNG, Janice K.
;  APPLICANT:  TANG, Y. Tom; AZIMZAI, Yalda
;  APPLICANT:  BURRILL, John D.; MARCUS, Gregory A.
;  APPLICANT:  ZINGLER, Kurt A.; LU, Dyung Aina M.
;  APPLICANT:  LAL, Preeti G.; RAMKUMAR, Jayalaxmi
;  APPLICANT:  WARREN, Bridget A.; KEARNEY, Liam
;  APPLICANT:  POLICKY, Jennifer L.; THANGAVELU, Kavitha
;  APPLICANT:  BURFORD, Neil
; TITLE OF INVENTION: HUMAN KINASES
; FILE REFERENCE: PF-0209 USN
; CURRENT APPLICATION NUMBER: US/10/362,892
; CURRENT FILING DATE:  2003-02-25
; PRIOR APPLICATION NUMBER: PCT/US01/27219
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/229,873
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 60/231,357
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: US 60/232,654
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US 60/234,902
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US 60/236,499
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: US 60/238,389
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/240,542
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PERL Program
; SEQ ID NO 8
;   LENGTH: 2429
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20040038881A1 079284CD1
US-10-362-892-8

  Query Match             3.3%;  Score 429.5;  DB 12;  Length 2429;
  Best Local Similarity   19.9%;  Pred. No. 4.2e-11;
  Matches  551;  Conservative  305;  Mismatches  982;  Indels  931;  Gaps  139;

Qy         38 GLLEYQHH-----SRDYASHLSPGSI-----IQPQRRRPSLLSEFQPGNERSQELHLRPE 87

Db        215 GVLSFTHHQIIELARDCLDKSHQGLITSRYFLELQHKLDKLL----------QEAHDRSE 264

Qy         88 SHSYLPELGKSEMEFIESKRPRLELLPDPLLRPSPLLATGQPAGSEDLTKDRSLTGKLEP 147

Db        265 S---------GELAFIKQLVRKILIV---IARPARLLEC-----------------LE- 293

Qy        148 VSPPSPPHTDPE-----LELVPPRLSKEELIQNMDRVDREITMVEQQISKLKKKQQQLEE 202

Db        294 --------FDPEEFYYLLEAAEGHAKEGQGIKT--DIPRYI------ISQLGLNKDPLEE 337

Qy        203 EA---------AKPPEPEKPVSPPPIESKHRSLVQIIYDENRKKAEAAHRILEGLGPQVE 253

Db        338 MAHLGNYDSGTAETPETDESVSSSNASLKLR----------RKPRE-------------- 373

Qy        254 LPLYNQPSDTRQYHENIKI--NQAMRKKLILYFKRRNHARKQWKQKFCQRY-----DQLM 306

Db        374 -------SD----FETIKLISNGAYG---AVYFVRHKESRQRFAMKKINKQNLILRNQIQ 419

Qy        307 EA-LEKKVERIENNPRRRAKESKVREYYEKQFPEIRKQRELQERMQSRVG------QRGS 359

Db        420 QAFVERDILTFAENP----------FVVSMYCSFETRRHLCMVMEYVEGGDCATLMKNM 468

Qy        360 G-LSMSAARSEHEVSEIIDGLSEQENLEKQMRQLAVIPPMLYDADQQRIKFINMN----G 414
```





```
Qy       2310 GTEIFNMPAITGTGLMTYRSQAVQEH----ASTNMGLEAIIRKALMGKYDQWEESPPLSA 2365
                                                                          
Db       2248 ----------SATGQSSFRSTALPEKSLSCSSSFPETRAGVREASAASSD---------- 2287

Qy       2366 NAFNPLNASASLPAAMPITAADGRSDHTLTSPGGGGK---AKVSGRPSSR---------- 2412
                                                                          
Db       2288 ------TSSAKAAGGMLELPAPSNRDHRKAQPAGEGRTHMTKSDSLPSFRVSTLPLESHH 2341

Qy       2413 ------------------------KAKSPAPGLASGDRPPSVSSVHSEGDCNRRTPLTN 2447
                                                                          
Db       2342 PDPNTMGGASHRDRALSVTATVGETKGKDPAPA-----QPPPARKQNVGRDVTKPSPAPN 2396

Qy       2448 RVWEDRPSS 2456
                                                                          
Db       2397 ---TDRPIS 2402
```

RESULT 26
US-10-288-798-8
; Sequence 8, Application US/10288798
; Publication No. US20030207299A1
; GENERAL INFORMATION:
; APPLICANT: BANDMAN, Olga; NGUYEN, Danniel B;
; APPLICANT: WALIA, Narinder K.; HAFALIA, April J.A.;
; APPLICANT: YAO, Monique G.; GANDHI, Ameena R.;
; APPLICANT: GURURAJAN, Rajagopal; DING, Li;
; APPLICANT: PATTERSON, Chandra; YUE, Henry;
; APPLICANT: BAUGHN, Mariah R.; TRIBOULEY, Catherine M.;
; APPLICANT: THORNTON, Michael; ELLIOTT, Vicki S.;
; APPLICANT: LU, Yan; ISON, Craig H.;
; APPLICANT: AU-YOUNG, Janice; TANG, Y. Tom;
; APPLICANT: AZIMZAI, Yalda; BURRILL, John D.;
; APPLICANT: MARCUS, Gregory A.; ZINGLER, Kurt A.;
; APPLICANT: LU, Dyung Aina M.; LAL, Preeti G.;
; APPLICANT: RAMKUMAR, Jayalaxmi; WARREN, Bridget A.;
; APPLICANT: KEARNEY, Liam; POLICKY, Jennifer L.;
; APPLICANT: THANGAVELU, Kavitha; BURFORD, Neil
; TITLE OF INVENTION: HUMAN KINASES
; FILE REFERENCE: PI-0209 USA
; CURRENT APPLICATION NUMBER: US/10/288,798
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: PCT/US01/27219
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/240,542
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/238,389
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/236,499
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: US 60/234,902
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US 60/232,654
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US 60/231,357
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: US 60/229,873
; PRIOR FILING DATE: 2000-08-31
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PERL Program
; SEQ ID NO 8
; LENGTH: 2429
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030207299A1 079284CD1
US-10-288-798-8

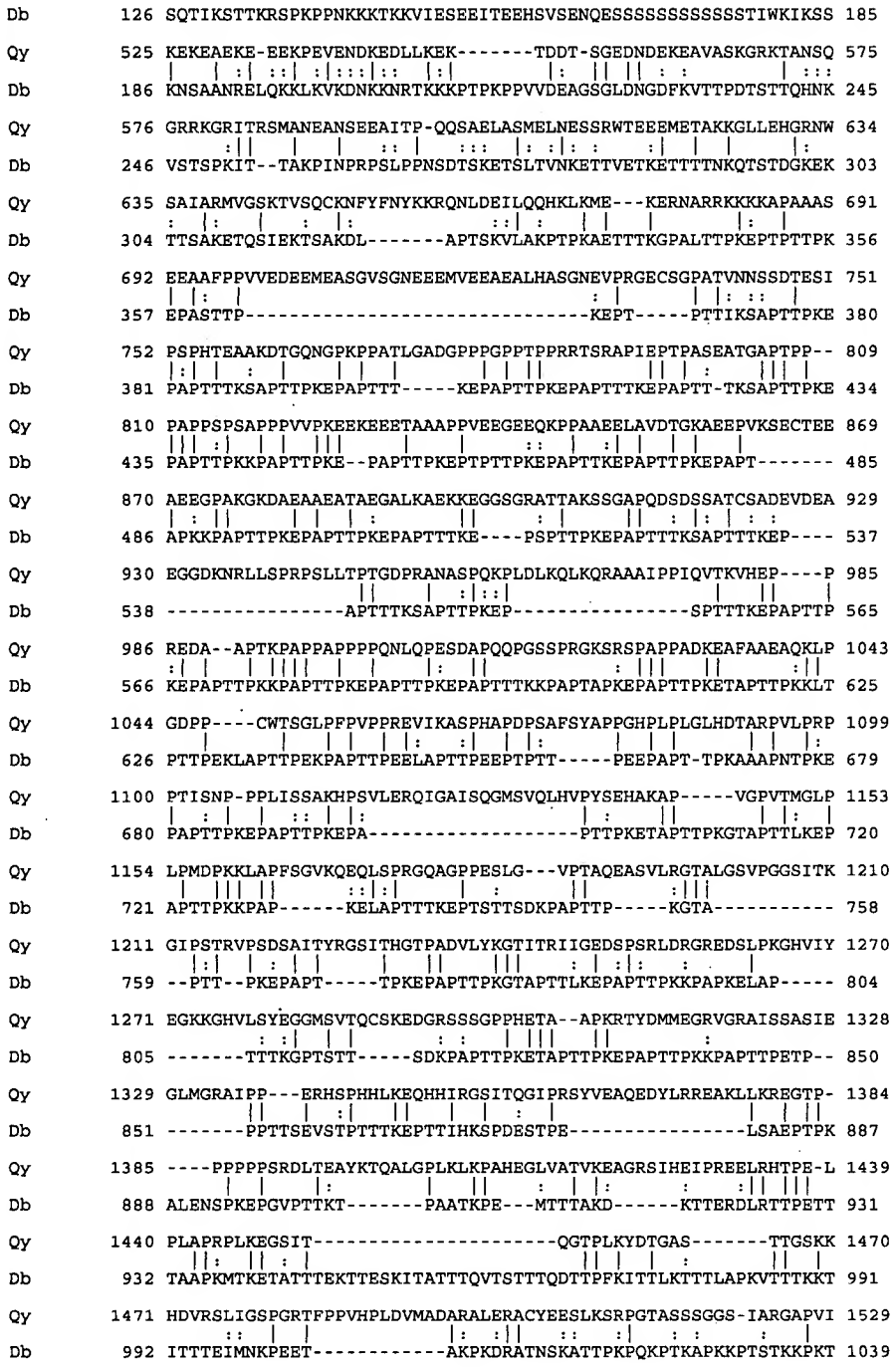
Query Match 3.3%; Score 429.5; DB 15; Length 2429;
Best Local Similarity 19.9%; Pred. No. 4.2e-11;
Matches 551; Conservative 305; Mismatches 982; Indels 931; Gaps 139;

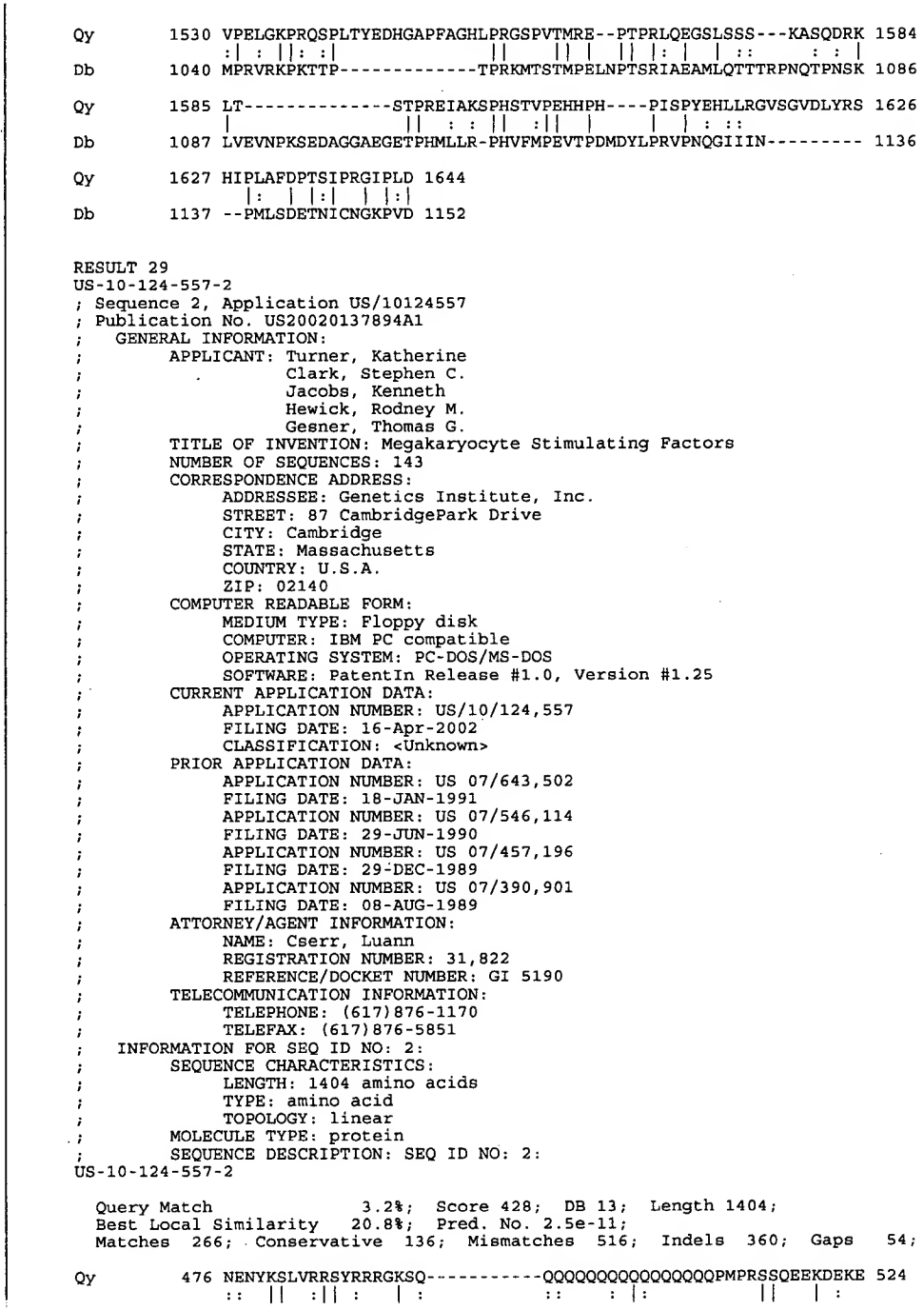
```
Qy         38 GLLEYQHH-----SRDYASHLSPGSI-----IQPQRRRPSLLSEFQPGNERSQELHLRPE 87
                                                                          
Db        215 GVLSFTHHQIIELARDCLDKSHQGLITSRYFLELQHKLDKLL----------QEAHDRSE 264

Qy         88 SHSYLPELGKSEMEFIESKRPRLELLPDPLLRPSPLLATGQPAGSEDLTKDRSLTGKLEP 147
                                                                          
Db        265 S---------GELAFIKQLVRKILIV---IARPARLLEC----------------LE- 293

Qy        148 VSPPSPPHTDPE-----LELVPPRLSKEELIQNMDRVDREITMVEQQISKLKKKQQQLEE 202
                                                                          
Db        294 --------FDPEEFYYLLEAAEGHAKEGQGIKT--DIPRYI------ISQLGLNKDPLEE 337

Qy        203 EA---------AKPPEPEKPVSPPPIESKHRSLVQIIYDENRKKAEAAHRILEGLGPQVE 253
                                                                          
Db        338 MAHLGNYDSGTAETPETDESVSSSNASLKLR----------RKPRE-------------- 373

Qy        254 LPLYNQPSDTRQYHENIKI--NQAMRKKLILYFKRRNHARKQWKQKFCQRY-----DQLM 306
                                                                          
Db        374 -------SD----FETIKLISNGAYG---AVYFVRHKESRQRFAMKKINKQNLILRNQIQ 419

Qy        307 EA-LEKKVERIENNPRRRAKESKVREYYEKQFPEIRKQRELQERMQSRVG------QRGS 359
                                                                          
Db        420 QAFVERDILTFAENP----------FVVSMYCSFETRRHLCMVMEYVEGGDCATLMKNM 468

Qy        360 G-LSMSAARSEHEVSEIIDGLSEQENLEKQMRQLAVIPPMLYDADQQRIKFINMN----G 414
                                                                          
Db        469 GPLPVDMARM--YFAETVLALEYLHNYGIVHRDLK--PDNLLVTSMGHIKLTDFGLSKVG 524

Qy        415 LMADPMKVYK----------DRQVMNMWSEQEKETFREKFMQHPKNFGLIASFLERKTV 463
                                                                          
Db        525 LMSMTTNLYEGHIEKDAREFLDKQVCGTPEYIAPEVILRQGYGKPVDWWAMGIILYEFLV 584

Qy        464 AECVLYYYLT----------------KKNE----NYKSLV-----RRSYRRRGKSQQQQQ 498
                                                                          
Db        585 G-CVPFFGDTPEELFGQVISDEINWPEKDEAPPPDAQDLITLLLRQNPLERLGTGGAYEV 643

Qy        499 QQQQQQQQQQQQPMPRSSQEEKDEKEKEKEAEKEEEKPEVENDKEDLLKEKTDDTSGED- 557
                                                                          
Db        644 KQHRFFRSLDWNSLLRQKAEFIPQLESEDDTSYFDTRSEKYHHME---TEEEDDTNDEDF 700

Qy        558 NDEKEAVAS---KGRKTANSQGRRKGRITRSMANEANSEEAITPQQSAELASME-LNESS 613
                                                                          
Db        701 NVEIRQFSSCSHRFSKVFSS----IDRITQNSAEE--KEDSVDKTKSTTLPSTETLSWSS 754

Qy        614 RWTE-EEMETAKKGLLEHGRNWSAIARMVGSKTVSQCKNFYFNYKKRQNLDEILQQHKLK 672
                                                                          
Db        755 EYSEMQQLSTSNSSDTESNR----------------------HKLSSGL---LPKLAIS 788

Qy        673 MEKERNARRKKKKAPAAASEEAAFPPVVEDEEMEASGVSGNEEEMVEEAEALHASGNEV- 731
                                                                          
Db        789 TEGEQD---EAASCPGDPHEEPGKPALPPEE------CAQEEPEVTTPASTISSSTLSVG 839

Qy        732 ----------PRGECSGPATVNNSSDTESIPSPH--------TEAAKDTGQNGPKPPATL 773
                                                                          
Db        840 SFSEHLDQINGRSECVD--STDNSSKPSSEPASHMARQRLESTEKKKISG----KVTKSL 893

Qy        774 GADG------------PPPGPPTPPRRTSRAPIEPTPASEATGAPTPPPAPPSPSAPPPV 821
                                                                          
Db        894 SASALSLMIPGDMFAVSPLGSPMSPHSLS-----SDPSSSRDSSPSRDSSAASASPHQPI 948

Qy        822 VPKEEKEEE--TAAAPPVEEGE------------EQKPPAAEELAVDTGKAEEPVKSE- 865
                                                                          
Db        949 VIHSSGKNYGFTIRAIRVYVGDSDIYTVHHIVWNVEEGSPAC-QAGLKAGDLITPINGEP 1007

Qy        866 -----CTEEAE----------------------EGPAKGKDAEAAEATAEGALKAEKKEG 898
                                                                          
Db       1008 VHGLVHTEVIELLLKSGNKVSITTTPFENTSIKTGPARRNSYKS--RMVRRSKKSKKKES 1065

Qy        899 GSGRAT----TAKSSGAPQDSDSSATCSADEVDEAEGGDKNRLLSPRPSLLTPTGDPRAN 954
                                                                          
Db       1066 LERRRSLFKKLAKQPSPLLHTSRSFSC-----------LNRSLSSGESL---PGSPTHS 1110

Qy        955 ASPQKPLDLKQLKQRAAAIPPIQVTKVHEPPREDAAPTKPAPPAPPPPQNLQPESDAPQQ 1014
                                                                          
```

```
Db       1111 LSPRSP----------------------TPSYRSTPDFPS--GTNSSQSSSPSSSAPNS 1145

Qy       1015 PGSSPRGKSRSP-----APPADKEAFAAEAQKLPGDPPCWTSGLPFPVPPREVIKASPHA 1069

Db       1146 PAGS--GHIRPSTLHGLAPKLGGQRYRSGRRKSAGNIP----------------LSPLA 1186

Qy       1070 --PDPSAFSYAPPGHPLPLGLHDTARPVLPR--PPTISNPPPLISSAKHPSVLERQIGAI 1125

Db       1187 RTPSPTPQPTSPQRSPSPLLGHSLGNSKIAQAFPSKMHSPPTIVR-------------- 1231

Qy       1126 SQGMSVQLHVPYSEHAKAPVGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPESLG 1185

Db       1232 --------HIVRPKSAEPPRSPL---LKRVQSEEKLSPSYGSDKKHLCSR-----KHSLE 1275

Qy       1186 VPTAQEASVLRGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHGTPADVLYKGTITR 1245

Db       1276 V-TQEE--VQREQSQREAPLQSLDENVCDVPPLSRARPVEQGCL----------KRPVSR 1322

Qy       1246 IIG-EDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPHET 1304

Db       1323 KVGRQESVDDLDR---DKL-KAKVVVK------------------KADGF--------- 1350

Qy       1305 AAPKRTYDMMEGRVGRAISSASIEGLMGRAIPPERHSPHHLKEQHHIRGSITQGIPRSYV 1364

Db       1351 -------------------------------PEKQESH---QKSHGPGSDLENFALFKL 1375

Qy       1365 EAQEDYLRREAKLLKREGTPPPPPPSRDLTEAYKTQA-------LGPLKLKPAHEGLVAT 1417

Db       1376 E------EREKKVY---------PKAVERSSTFENKASMQEAPPLGSLLKDALHKQASVR 1420

Qy       1418 VKEAGRSIHEIPREELRHTPELPLAPRPLKEGSITQGTPLKYDTGASTTG---SKKHDVR 1474

Db       1421 ASEGAMSDGPVPAEHRQGGGDFRRAPAP---GTLQDGLCHSLDRGISGKGEGTEKSSQAK 1477

Qy       1475 SLIGSPGRTFPPVHPLDVMADARALERACYEESLKSRPGTASSSGGSIARGAPVIVPELG 1534

Db       1478 ELLRCE-KLDSKLANIDYLRKKMSLEDK--EDNL----------------CPVL----- 1512

Qy       1535 KPRQSPLTYEDHGAPFAGHLPRGSPVT----MREPTPRLQEGS-LSSSKASQDRKLTSTP 1589

Db       1513 KPKMTAGSHE--------CLP-GNPVRPTGGQQEPPPASESRAFVSSTHAAQMSAVSFVP 1563

Qy       1590 REI--------AKSPHSTVPEHHPHPISPYEHLLRG----------------VSGVDLY 1624

Db       1564 LKALTGRVDSGTEKPGLVAPE-SPVRKSPSEYKLEGRSVSCLKPIEGTLDIALLSGPQAS 1622

Qy       1625 RSHIP---LAFDPT-------SIPRGIPL------DAAAAYYL-PRHLAPNPTYPHLYPP 1667

Db       1623 KTELPSPESAQSPSPSGDVRASVPPVLPSSSGKKNDTTSARELSPSSLKMNKSY--LLEP 1680

Qy       1668 YLIRGYPDTAALENRQTII---NDYITSQQMHHNTA----TAMAQRADMLRGLSPRESSL 1720

Db       1681 WFL---PPSRGLQNSPAVSLPDPEFKRDRKGPHPTARSPGTVMESNPQQREGSSPKHQ-- 1735

Qy       1721 ALNYAAGPRGIIDLSQVPHLPVLVPPTPGTPATAMDRLAYLPTAPQPFSSRHSSSPLSPG 1780

Db       1736 --DHTTDPKLLTCLGQNLHSPDLARP--------------------------RCPLPPE 1766

Qy       1781 GPTHLTKPTTTSSSER-------ERDRDR--------------ERDRDREREKSIL---- 1815

Db       1767 ASPSREKPGLRESSERGPPTARSERSAARADTCREPSMELCFPETAKTSDNSKNLLSVGR 1826

Qy       1816 ------TSTTTVEHAPIWRPGTEQSSGSSGSSGGGGGSSSRPASHSHAHQHSPISPRTQD 1869

Db       1827 THPDFYTQTQAMEKA--WAPG-----GKTNHKDGPG--EARPPPRDNSSLHSAGIPCEKE 1877

Qy       1870 ALQQRPSVLHNTGMKGIITAVEPSKPTVLRSTSTSSPVRPAATFPPATHCPLGGTLDGVY 1929

Db       1878 LGKVR-------------RGVEPKPEALLARRSLQ---------PPGIESEKSEKLSS-F 1914

Qy       1930 PTLM-----EPVLLPKEAPRVARPER-PRADTGHAFLAKPPAR---------SGLEPASS 1974

Db       1915 PSLQKDGAKEPE--RKEQPLQRHPSSIPPPPLTAKDLSSPAARQHCSSPSHASGREPGAK 1972

Qy       1975 PS---KGSEPRPLVPPVSGHATIARTPAKNLAPHHASPDPPAPPASASDPHREKTQSKPF 2031

Db       1973 PSTAEPSSSPQDPPKPVAAHS--------ESSSHKPRPGPDPGPPKTKHPDRSLSSQKP- 2023

Qy       2032 SI-----QELELRSLG---YHGSSYSPEG--VEPVSPVS----SPSLTHDKGLPKHLEEL 2077

Db       2024 SVGATKGKEPATQSLGGSSREGKGHSKSGPDVFPATPGSQNKASDGIGQGEGGPSVPLHT 2083

Qy       2078 DKSHLEGELRPKQPG-PVKLGGEAAHLPHLRPLPESQPSSSPLLQTAPGVKGHQRVVTLA 2136

Db       2084 DRAPLDAKPQPTSGGRPLEVLEKPVHLPRPGHPGPSEPADQKL--SAVGEK--------- 2132

Qy       2137 QHISEVITQDYTRHHPQQLSAPLPAPLYSFPGASCPVLDLR---RPPSDLYLPPPDHGAP 2193

Db       2133 --------QTLSPKHPK------PSTVKDCP-TLCKQTDNRQTDKSPSQ----------P 2167

Qy       2194 ARGSPHSEGGKRSPE----PNKTSVLGGGEDGIEPVSPPEGMTEPGHSRSAVYPLLYRDG 2249

Db       2168 AANTDRRAEGKKCTEALYAPAEGDKLEAGLSFVHSENRLKGAERPAAGVGKGFP------ 2221

Qy       2250 EQTEPSRMGSKSPGNTSQPPAFFSKLTESNSAMVKSKKQEINKKLNTHNRNEPEYNISQP 2309

Db       2222 ------EARGKGPG-PQKPP------TEAD------------KPNGMKRSP-------- 2247

Qy       2310 GTEIFNMPAITGTGLMTYRSQAVQEH----ASTNMGLEAIIRKALMGKYDQWEESPPLSA 2365

Db       2248 ----------SATGQSSFRSTALPEKSLSCSSSFPETRAGVREASAASSD---------- 2287

Qy       2366 NAFNPLNASASLPAAMPITAADGRSDHTLTSPGGGGK---AKVSGRPSSR---------- 2412

Db       2288 ------TSSAKAAGGMLELPAPSNRDHRKAQPAGEGRTHMTKSDSLPSFRVSTLPLESHH 2341

Qy       2413 ------------------------KAKSPAPGLASGDRPPSVSSVHSEGDCNRRTPLTN 2447

Db       2342 PDPNTMGGASHRDRALSVTATVGETKGKDPAPA-----QPPPARKQNVGRDVTKPSPAPN 2396

Qy       2448 RVWEDRPSS 2456

Db       2397 ---TDRPIS 2402
```

RESULT 27
US-10-124-557-52
; Sequence 52, Application US/10124557
; Publication No. US20020137894A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991

```
;             APPLICATION NUMBER: US 07/546,114
;             FILING DATE: 29-JUN-1990
;             APPLICATION NUMBER: US 07/457,196
;             FILING DATE: 29-DEC-1989
;             APPLICATION NUMBER: US 07/390,901
;             FILING DATE: 08-AUG-1989
;       ATTORNEY/AGENT INFORMATION:
;             NAME: Cserr, Luann
;             REGISTRATION NUMBER: 31,822
;             REFERENCE/DOCKET NUMBER: GI 5190
;       TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (617)876-1170
;             TELEFAX: (617)876-5851
;   INFORMATION FOR SEQ ID NO: 52:
;       SEQUENCE CHARACTERISTICS:
;             LENGTH: 1363 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 52:
US-10-124-557-52

  Query Match             3.2%;  Score 428;  DB 13;  Length 1363;
  Best Local Similarity   20.8%;  Pred. No. 2.5e-11;
  Matches  266;  Conservative  136;  Mismatches  516;  Indels  360;  Gaps   54;

Qy        476 NENYKSLVRRSYRRRGKSQ----------QQQQQQQQQQQQQQQQPMPRSSQEEKDEKE 524
              ::  ||  :|| :   | :          ::    : |:          ||   | :
Db         85 SQTIKSTTKRSPKPPNKKKTKKVIESEEITEEHSVSENQESSSSSSSSSSSTIWKIKSS 144

Qy        525 KEKEAEKE-EEKPEVENDKEDLLKEK-------TDDT-SGEDNDEKEAVASKGRKTANSQ 575
              |   | :| ::| :|:::|::  |:|        |:  || || : :         | :::
Db        145 KNSAANRELQKKLKVKDNKKNRTKKKPTPKPPVVDEAGSGLDNGDFKVTTPDTSTTQHNK 204

Qy        576 GRRKGRITRSMANEANSEEAITP-QQSAELASMELNESSRWTEEEMETAKKGLLEHGRNW 634
                 :||   |   |   :: |   :::  |: :|: :    :|  |  |      |:
Db        205 VSTSPKIT--TAKPINPRPSLPPNSDTSKETSLTVNKETTVETKETTTTNKQTSTDGKEK 262

Qy        635 SAIARMVGSKTVSQCKNFYFNYKKRQNLDEILQQHKLKME---KERNARRKKKKAPAAAS 691
              :  |:   |   :  |:          ::| :   | |   |       |:  |
Db        263 TTSAKETQSIEKTSAKDL-------APTSKVLAKPTPKAETTTKGPALTTPKEPTPTTPK 315

Qy        692 EEAAFPPVVEDEEMEASGVSGNEEEMVEEAEALHASGNEVPRGECSGPATVNNSSDTESI 751
              | |:  |                              : |      | |: ::  |
Db        316 EPASTTP------------------------------KEPT-----PTTIKSAPTTPKE 339

Qy        752 PSPHTEAAKDTGQNGPKPPATLGADGPPPGPPTPPRRTSRAPIEPTPASEATGAPTPP-- 809
              |:| |  :  |    | | |       | | ||          || | :    ||| |
Db        340 PAPTTTKSAPTTPKEPAPTTT-----KEPAPTTPKEPAPTTTKEPAPTT-TKSAPTTPKE 393

Qy        810 PAPPSPSAPPPVVPKEEKEEETAAAPPVEEGEEQKPPAAEELAVDTGKAEEPVKSECTEE 869
              ||| :|  | |  |||     |    |      ::   |  :| |  | | |
Db        394 PAPTTPKKPAPTTPKE--PAPTTPKEPTPTTPKEPAPTTKEPAPTTPKEPAPT------- 444

Qy        870 AEEGPAKGKDAEAAEATAEGALKAEKKEGGSGRATTAKSSGAPQDSDSSATCSADEVDEA 929
              | : ||      | |  :       ||      : |     ||  : |: | : :
Db        445 APKKPAPTTPKEPAPTTPKEPAPTTTKE----PSPTTPKEPAPTTTKSAPTTTKEP---- 496

Qy        930 EGGDKNRLLSPRPSLLTPTGDPRANASPQKPLDLKQLKQRAAAIPPIQVTKVHEP----P 985
                          ||     |  :|::|              |   ||     |
Db        497 ----------------APTTTKSAPTTPKEP---------------SPTTTKEPAPTTP 524

Qy        986 REDA--APTKPAPPAPPPPQNLQPESDAPQQPGSSPRGKSRSPAPPADKEAFAAEAQKLP 1043
              :| |   | ||||  | |    |:  ||            : |||   ||      :||
Db        525 KEPAPTTPKKPAPTTPKEPAPTTPKEPAPTTTKKPAPTAPKEPAPTTPKETAPTTPKKLT 584

Qy       1044 GDPP----CWTSGLPFPVPPREVIKASPHAPDPSAFSYAPPGHPLPLGLHDTARPVLPRP 1099
                |       |  | |  | |:   :|  | |:      |  | |      | |  |:
Db        585 PTTPEKLAPTTPEKPAPTTPEELAPTTPEEPTPTT-----PEEPAPT-TPKAAAPNTPKE 638

Qy       1100 PTISNP-PPLISSAKHPSVLERQIGAISQGMSVQLHVPYSEHAKAP-----VGPVTMGLP 1153
              |  : |  |  :: | |:                | :    ||       | |:  |
Db        639 PAPTTPKEPAPTTPKEPA------------------PTTPKETAPTTPKGTAPTTLKEP 679

Qy       1154 LPMDPKKLAPFSGVKQEQLSPRGQAGPPESLG---VPTAQEASVLRGTALGSVPGGSITK 1210
                |  ||| ||      ::|:|     |  :      ||       :|||
Db        680 APTTPKKPAP------KELAPTTTKEPTSTTSDKPAPTTP-----KGTA---------- 717

Qy       1211 GIPSTRVPSDSAITYRGSITHGTPADVLYKGTITRIIGEDSPSRLDRGREDSLPKGHVIY 1270
                |:|  | : | |      |   ||     |||     : | :|:   :     |
Db        718 --PTT--PKEPAPT-----TPKEPAPTTPKGTAPTTLKEPAPTTPKKPAPKELAP----- 763

Qy       1271 EGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPHETA--APKRTYDMMEGRVGRAISSASIE 1328
                    : :|  | |        : :   | |||    ||          :
Db        764 -------TTTKGPTSTT-----SDKPAPTTPKETAPTTPKEPAPTTPKKPAPTTPETP-- 809

Qy       1329 GLMGRAIPP---ERHSPHHLKEQHHIRGSITQGIPRSYVEAQEDYLRREAKLLKREGTP- 1384
                    ||   |  :|   ||   |  |  : |                  |  | ||
Db        810 -------PPTTSEVSTPTTTKEPTTIHKSPDESTPE---------------LSAEPTPK 846

Qy       1385 ----PPPPPSRDLTEAYKTQALGPLKLKPAHEGLVATVKEAGRSIHEIPREELRHTPE-L 1439
                   |  |    |:          |   ||    :  | |:       :    :|| |||
Db        847 ALENSPKEPGVPTTKT-------PAATKPE---MTTTAKD------KTTERDLRTTPETT 890

Qy       1440 PLAPRPLKEGSIT----------------------QGTPLKYDTGAS-------TTGSKK 1470
                 ||:  || : |                      || |  |  :        ||  |
Db        891 TAAPKMTKETATTTEKTTESKITATTTQVTSTTTQDTTPFKITTLKTTTLAPKVTTTKKT 950

Qy       1471 HDVRSLIGSPGRTFPPVHPLDVMADARALERACYEESLKSRPGTASSSGGS-IARGAPVI 1529
                  ::  |  |            |: :||   ::   :|   : :      :   |
Db        951 ITTTEIMNKPEET-----------AKPKDRATNSKATTPKPQKPTKAPKKPTSTKKPKT 998

Qy       1530 VPELGKPRQSPLTYEDHGAPFAGHLPRGSPVTMRE--PTPRLQEGSLSSS---KASQDRK 1584
                 :| : ||: :|            ||    || |  || |:|  |  | ::     : |
Db        999 MPRVRKPKTTP------------TPRKMTSTMPELNPTSRIAEAMLQTTTRPNQTPNSK 1045

Qy       1585 LT--------------STPREIAKSPHSTVPEHHPH----PISPYEHLLRGVSGVDLYRS 1626
                 |              ||  : : ||  :||  |      |  | : ::
Db       1046 LVEVNPKSEDAGGAEGETPHMLLR-PHVFMPEVTPDMDYLPRVPNQGIIIN--------- 1095

Qy       1627 HIPLAFDPTSIPRGIPLD 1644
                |:  | |:|  | |:|
Db       1096 --PMLSDETNICNGKPVD 1111


RESULT 28
US-09-802-207-30
; Sequence 30, Application US/09802207
; Publication No. US20020086824A1
; GENERAL INFORMATION:
;  APPLICANT: Warman, Matthew
;  APPLICANT:  Carpten, John
;  APPLICANT:  Trent, Jeffrey
;  APPLICANT:  Marcelino, Jose
;  TITLE OF INVENTION: Novel Methods and Reagents for the Treatment of Osteoarthritis
;  FILE REFERENCE: Case-06212
;  CURRENT APPLICATION NUMBER: US/09/802,207
;  CURRENT FILING DATE:  2001-08-29
;  PRIOR APPLICATION NUMBER: 09/619,175
;  PRIOR FILING DATE: 2000-07-19
;  PRIOR APPLICATION NUMBER: 60/145,328
;  PRIOR FILING DATE: 1999-07-23
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 30
;   LENGTH: 1404
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-802-207-30

  Query Match             3.2%;  Score 428;  DB 12;  Length 1404;
  Best Local Similarity   20.8%;  Pred. No. 2.5e-11;
  Matches  266;  Conservative  136;  Mismatches  516;  Indels  360;  Gaps   54;

Qy        476 NENYKSLVRRSYRRRGKSQ----------QQQQQQQQQQQQQQQQPMPRSSQEEKDEKE 524
              ::  ||  :|| :   | :          ::    : |:          ||   | :
```

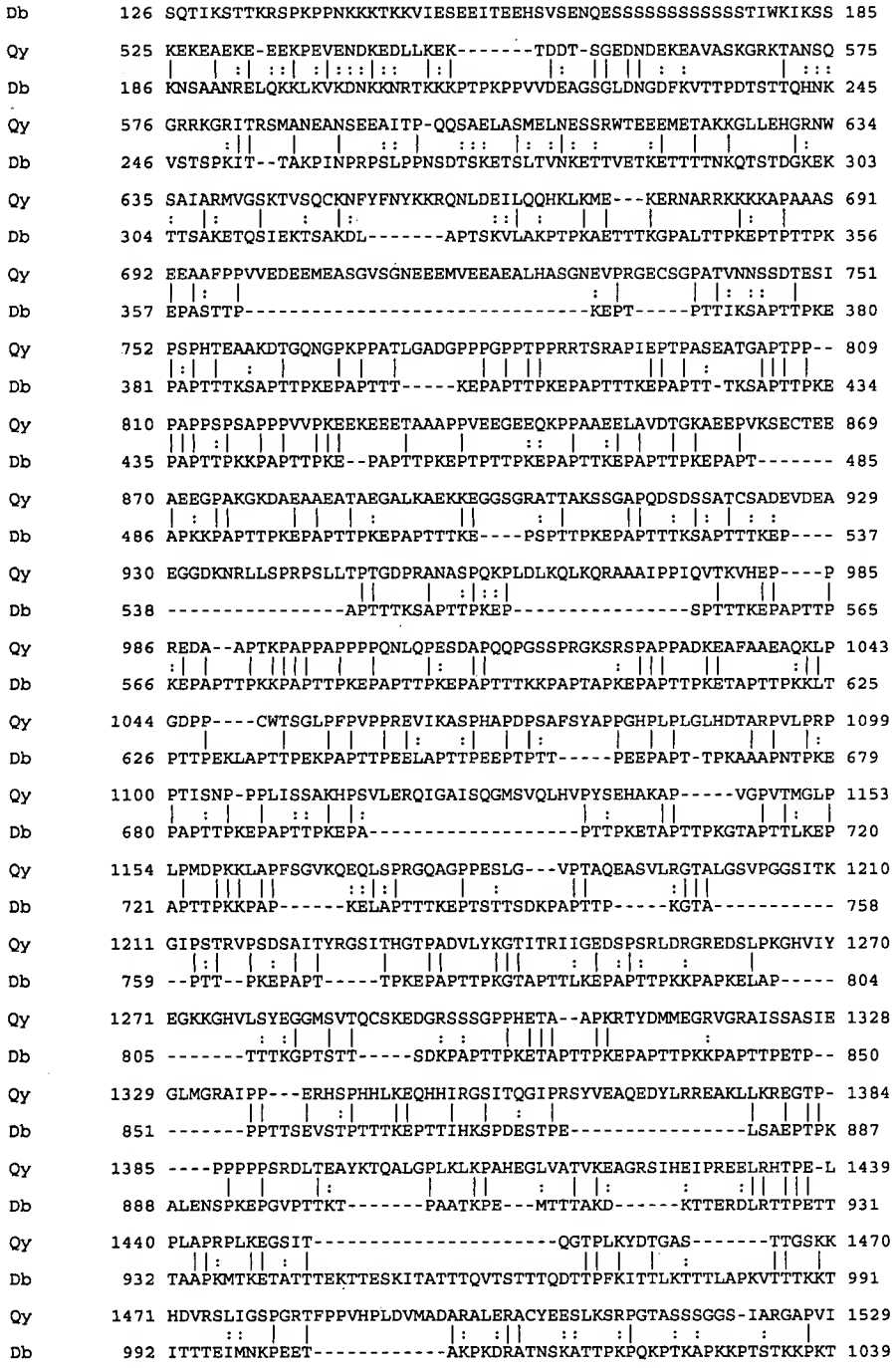


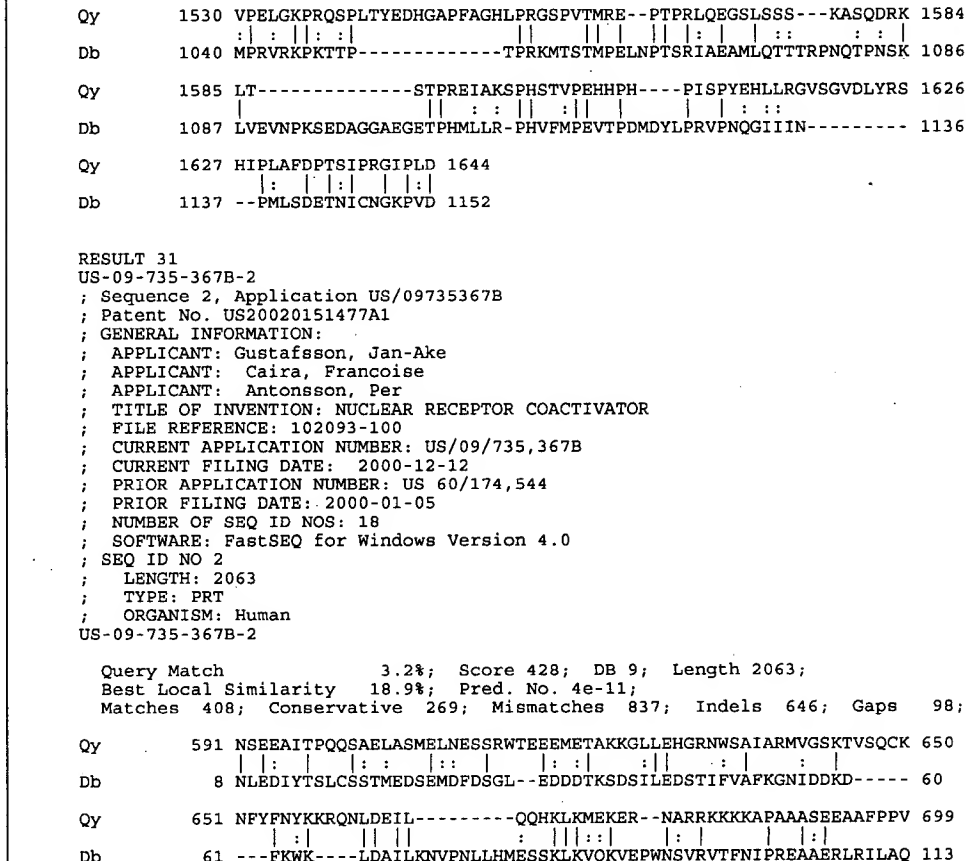


RESULT 29
US-10-124-557-2
; Sequence 2, Application US/10124557
; Publication No. US20020137894A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cserr, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)876-1170
; TELEFAX: (617)876-5851
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1404 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-124-557-2

Query Match 3.2%; Score 428; DB 13; Length 1404;
Best Local Similarity 20.8%; Pred. No. 2.5e-11;
Matches 266; Conservative 136; Mismatches 516; Indels 360; Gaps 54;

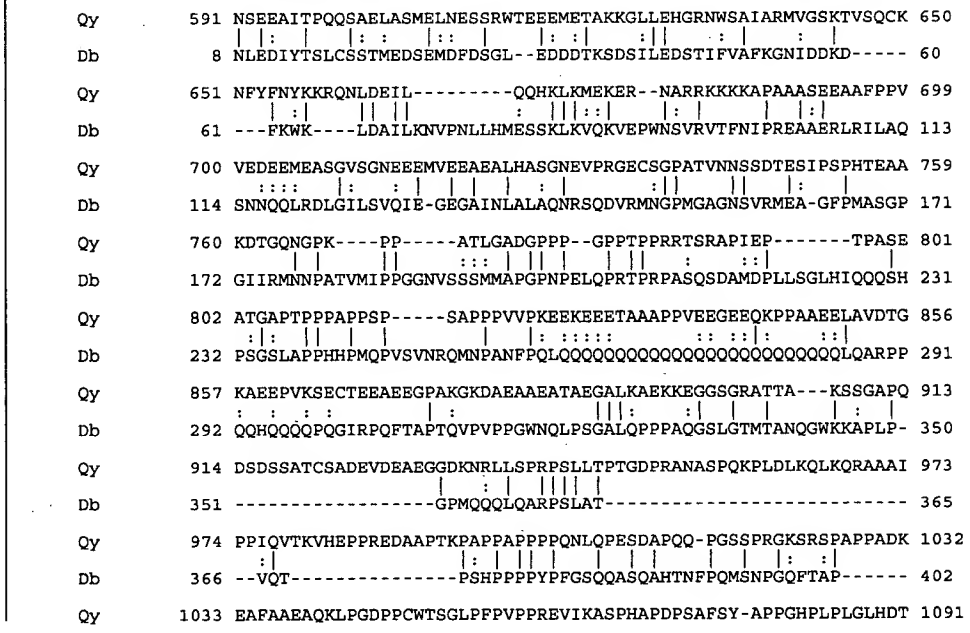
```
Db        126 SQTIKSTTKRSPKPPNKKKTKKVIESEEITEEHSVSENQESSSSSSSSSSSSTIWKIKSS 185

Qy        525 KEKEAEKE-EEKPEVENDKEDLLKEK-------TDDT-SGEDNDEKEAVASKGRKTANSQ 575

Db        186 KNSAANRELQKKLKVKDNKKNRTKKKPTPKPPVVDEAGSGLDNGDFKVTTPDTSTTQHNK 245

Qy        576 GRRKGRITRSMANEANSEEAITP-QQSAELASMELNESSRWTEEEMETAKKGLLEHGRNW 634

Db        246 VSTSPKIT--TAKPINPRPSLPPNSDTSKETSLTVNKETTVETKETTTTNKQTSTDGKEK 303

Qy        635 SAIARMVGSKTVSQCKNFYFNYKKRQNLDEILQQHKLKME---KERNARRKKKKAPAAAS 691

Db        304 TTSAKETQSIEKTSAKDL-------APTSKVLAKPTPKAETTTKGPALTTPKEPTPTTPK 356

Qy        692 EEAAFPPVVEDEEMEASGVSGNEEEMVEEAEALHASGNEVPRGECSGPATVNNSSDTESI 751

Db        357 EPASTTP-------------------------------KEPT-----PTTIKSAPTTPKE 380

Qy        752 PSPHTEAAKDTGQNGPKPPATLGADGPPPGPPTPPRRTSRAPIEPTPASEATGAPTPP-- 809

Db        381 PAPTTTKSAPTTPKEPAPTTT-----KEPAPTTPKEPAPTTTKEPAPTT-TKSAPTTPKE 434

Qy        810 PAPPSPSAPPPVVPKEEKEEETAAAPPVEEGEEQKPPAAEELAVDTGKAEEPVKSECTEE 869

Db        435 PAPTTPKKPAPTTPKE--PAPTTPKEPTPTTPKEPAPTTKEPAPTTPKEPAPT------- 485

Qy        870 AEEGPAKGKDAEAAEATAEGALKAEKKEGGSGRATTAKSSGAPQDSDSSATCSADEVDEA 929

Db        486 APKKPAPTTPKEPAPTTPKEPAPTTTKE----PSPTTPKEPAPTTTKSAPTTTKEP---- 537

Qy        930 EGGDKNRLLSPRPSLLTPTGDPRANASPQKPLDLKQLKQRAAAIPPIQVTKVHEP----P 985

Db        538 ----------------APTTTKSAPTTPKEP---------------SPTTTKEPAPTTP 565

Qy        986 REDA--APTKPAPPAPPPPQNLQPESDAPQQPGSSPRGKSRSPAPPADKEAFAAEAQKLP 1043

Db        566 KEPAPTTPKKPAPTTPKEPAPTTPKEPAPTTTKKPAPTAPKEPAPTTPKETAPTTPKKLT 625

Qy       1044 GDPP----CWTSGLPFPVPPREVIKASPHAPDPSAFSYAPPGHPLPLGLHDTARPVLPRP 1099

Db        626 PTTPEKLAPTTPEKPAPTTPEELAPTTPEEPTPTT-----PEEPAPT-TPKAAAPNTPKE 679

Qy       1100 PTISNP-PPLISSAKHPSVLERQIGAISQGMSVQLHVPYSEHAKAP-----VGPVTMGLP 1153

Db        680 PAPTTPKEPAPTTPKEPA------------------PTTPKETAPTTPKGTAPTTLKEP 720

Qy       1154 LPMDPKKLAPFSGVKQEQLSPRGQAGPPESLG---VPTAQEASVLRGTALGSVPGGSITK 1210

Db        721 APTTPKKPAP------KELAPTTTKEPTSTTSDKPAPTTP-----KGTA----------- 758

Qy       1211 GIPSTRVPSDSAITYRGSITHGTPADVLYKGTITRIIGEDSPSRLDRGREDSLPKGHVIY 1270

Db        759 --PTT--PKEPAPT-----TPKEPAPTTPKGTAPTTLKEPAPTTPKKPAPKELAP----- 804

Qy       1271 EGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPHETA--APKRTYDMMEGRVGRAISSASIE 1328

Db        805 -------TTTKGPTSTT-----SDKPAPTTPKETAPTTPKEPAPTTPKKPAPTTPETP-- 850

Qy       1329 GLMGRAIPP---ERHSPHHLKEQHHIRGSITQGIPRSYVEAQEDYLRREAKLLKREGTP- 1384

Db        851 -------PPTTSEVSTPTTTKEPTTIHKSPDESTPE---------------LSAEPTPK 887

Qy       1385 ----PPPPPSRDLTEAYKTQALGPLKLKPAHEGLVATVKEAGRSIHEIPREELRHTPE-L 1439

Db        888 ALENSPKEPGVPTTKT-------PAATKPE---MTTTAKD------KTTERDLRTTPETT 931

Qy       1440 PLAPRPLKEGSIT----------------------QGTPLKYDTGAS-------TTGSKK 1470

Db        932 TAAPKMTKETATTTEKTTESKITATTTQVTSTTTQDTTPFKITTLKTTTLAPKVTTTKKT 991

Qy       1471 HDVRSLIGSPGRTFPPVHPLDVMADARALERACYEESLKSRPGTASSSGGS-IARGAPVI 1529

Db        992 ITTTEIMNKPEET------------AKPKDRATNSKATTPKPQKPTKAPKKPTSTKKPKT 1039

Qy       1530 VPELGKPRQSPLTYEDHGAPFAGHLPRGSPVTMRE--PTPRLQEGSLSSS---KASQDRK 1584

Db       1040 MPRVRKPKTTP------------TPRKMTSTMPELNPTSRIAEAMLQTTTRPNQTPNSK 1086

Qy       1585 LT--------------STPREIAKSPHSTVPEHHPH----PISPYEHLLRGVSGVDLYRS 1626

Db       1087 LVEVNPKSEDAGGAEGETPHMLLR-PHVFMPEVTPDMDYLPRVPNQGIIIN--------- 1136

Qy       1627 HIPLAFDPTSIPRGIPLD 1644

Db       1137 --PMLSDETNICNGKPVD 1152


RESULT 30
US-10-124-557-62
; Sequence 62, Application US/10124557
; Publication No. US20020137894A1
;   GENERAL INFORMATION:
;        APPLICANT: Turner, Katherine
;                   Clark, Stephen C.
;                   Jacobs, Kenneth
;                   Hewick, Rodney M.
;                   Gesner, Thomas G.
;        TITLE OF INVENTION: Megakaryocyte Stimulating Factors
;        NUMBER OF SEQUENCES: 143
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Genetics Institute, Inc.
;             STREET: 87 CambridgePark Drive
;             CITY: Cambridge
;             STATE: Massachusetts
;             COUNTRY: U.S.A.
;             ZIP: 02140
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/124,557
;             FILING DATE: 16-Apr-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US 07/643,502
;             FILING DATE: 18-JAN-1991
;             APPLICATION NUMBER: US 07/546,114
;             FILING DATE: 29-JUN-1990
;             APPLICATION NUMBER: US 07/457,196
;             FILING DATE: 29-DEC-1989
;             APPLICATION NUMBER: US 07/390,901
;             FILING DATE: 08-AUG-1989
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Cserr, Luann
;             REGISTRATION NUMBER: 31,822
;             REFERENCE/DOCKET NUMBER: GI 5190
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (617)876-1170
;             TELEFAX: (617)876-5851
;   INFORMATION FOR SEQ ID NO: 62:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1404 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 62:
US-10-124-557-62

  Query Match.            3.2%;  Score 428;  DB 13;  Length 1404;
  Best Local Similarity   20.8%;  Pred. No. 2.5e-11;
  Matches  266;  Conservative  136;  Mismatches  516;  Indels  360;  Gaps   54;

Qy        476 NENYKSLVRRSYRRRGKSQ----------QQQQQQQQQQQQQQQQPMPRSSQEEKDEKE 524
```

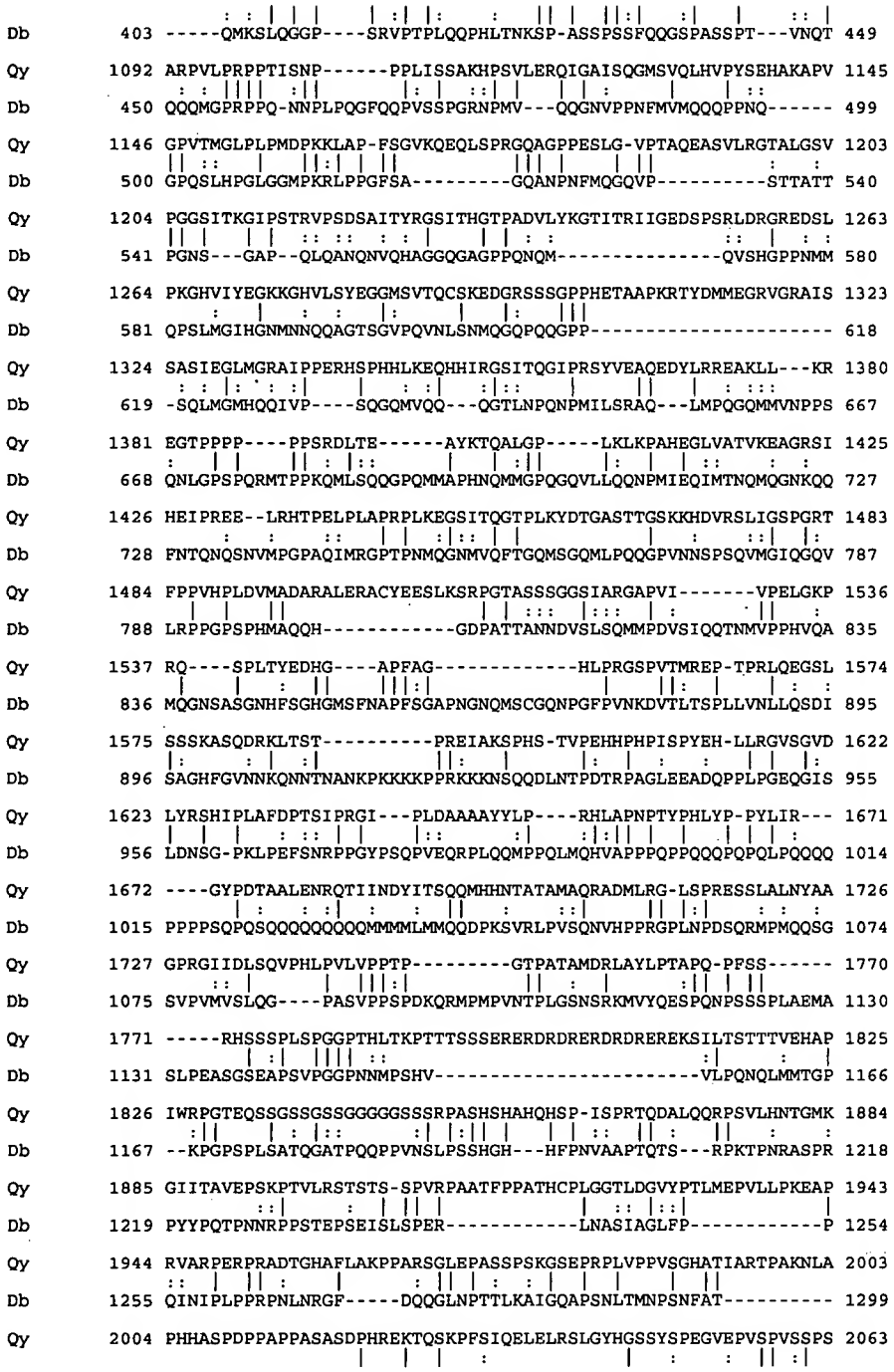


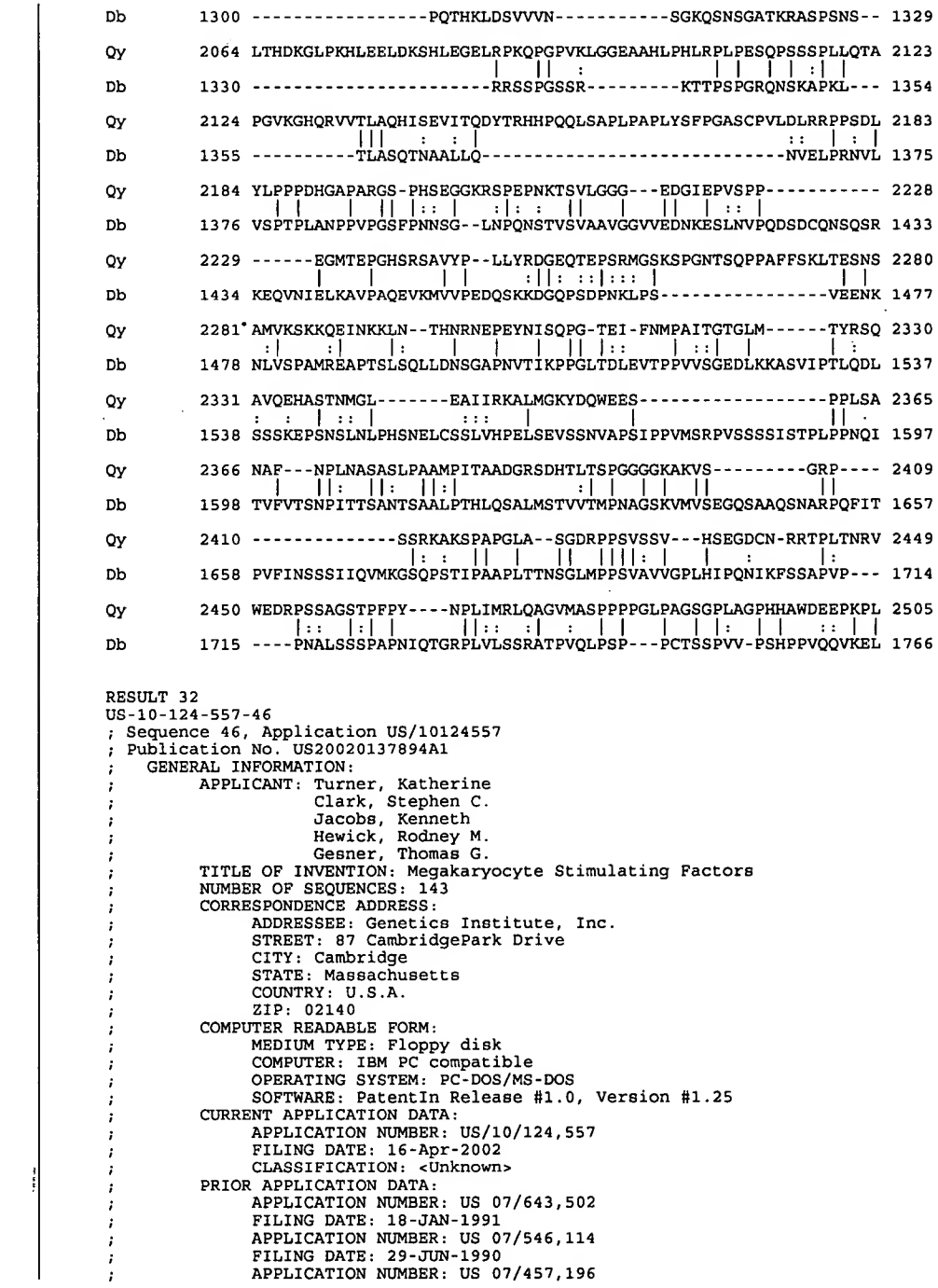


RESULT 31
US-09-735-367B-2
; Sequence 2, Application US/09735367B
; Patent No. US20020151477A1
; GENERAL INFORMATION:
; APPLICANT: Gustafsson, Jan-Ake
; APPLICANT: Caira, Francoise
; APPLICANT: Antonsson, Per
; TITLE OF INVENTION: NUCLEAR RECEPTOR COACTIVATOR
; FILE REFERENCE: 102093-100
; CURRENT APPLICATION NUMBER: US/09/735,367B
; CURRENT FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 60/174,544
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 2063
; TYPE: PRT
; ORGANISM: Human
US-09-735-367B-2

Query Match 3.2%; Score 428; DB 9; Length 2063;
Best Local Similarity 18.9%; Pred. No. 4e-11;
Matches 408; Conservative 269; Mismatches 837; Indels 646; Gaps 98;







RESULT 32
US-10-124-557-46
; Sequence 46, Application US/10124557
; Publication No. US20020137894A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; APPLICATION NUMBER: US 07/457,196

```
;              FILING DATE: 29-DEC-1989
;              APPLICATION NUMBER: US 07/390,901
;              FILING DATE: 08-AUG-1989
;         ATTORNEY/AGENT INFORMATION:
;              NAME: Cserr, Luann
;              REGISTRATION NUMBER: 31,822
;              REFERENCE/DOCKET NUMBER: GI 5190
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: (617)876-1170
;              TELEFAX: (617)876-5851
;    INFORMATION FOR SEQ ID NO: 46:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 1320 amino acids
;              TYPE: amino acid
;              TOPOLOGY: linear
;         MOLECULE TYPE: protein
;         SEQUENCE DESCRIPTION: SEQ ID NO: 46:
US-10-124-557-46

  Query Match             3.2%;  Score 425.5;  DB 13;  Length 1320;
  Best Local Similarity   21.2%;  Pred. No. 3.1e-11;
  Matches  261;  Conservative  127;  Mismatches  491;  Indels  355;  Gaps   53;

Qy        508 QQQPMPRSSQEEKDEKEKEKEAEKEEEKPEVENDKEDLLKEK-------TDDT-SGEDND 559
              :: | |         |:| |:  : ||  ||:::|::  |:|         |:  || ||
Db         93 KRSPKP-------PNKKKTKKVIESEEITEVKDNKKNRTKKKPTPKPPVVDEAGSGLDNG 145

Qy        560 EKEAVASKGRKTANSQGRRKGRITRSMANEANSEEAITP-QQSAELASMELNESSRWTEE 618
              : :        | :::     :||   |    |   :: |   :::  |: :|: :    :
Db        146 DFKVTTPDTSTTQHNKVSTSPKIT--TAKPINPRPSLPPNSDTSKETSLTVNKETTVETK 203

Qy        619 EMETAKKGLLEHGRNWSAIARMVGSKTVSQCKNFYFNYKKRQNLDEILQQHKLKME---K 675
              |  |  |      |:  :  |:   |   :  |:            ::| :   | |   |
Db        204 ETTTTNKQTSTDGKEKTTSAKETQSIEKTSAKDL-------APTSKVLAKPTPKAETTTK 256

Qy        676 ERNARRKKKKAPAAASEEAAFPPVVEDEEMEASGVSGNEEEMVEEAEALHASGNEVPRGE 735
                    |:  |    | |:  |                              : |
Db        257 GPALTTPKEPTPTTPKEPASTTP------------------------------KEPT-- 283

Qy        736 CSGPATVNNSSDTESIPSPHTEAAKDTGQNGPKPPATLGADGPPPGPPTPPRRTSRAPIE 795
                 | |: ::  |   |:| |  :  |   | |  |       | | ||         |
Db        284 ---PTTIKSAPTTPKEPAPTTTKSAPTTPKEPAPTTT-----KEPAPTTPKEPAPTTTKE 335

Qy        796 PTPASEATGAPTPP--PAPPSPSAPPPVVPKEEKEEETAAAPPVEEGEEQKPPAAEELAV 853
              | | :    ||| |  ||| :|  | |  |||      |     |      ::  |  :| |
Db        336 PAPTT-TKSAPTTPKEPAPTTPKKPAPTTPKE--PAPTTPKEPTPTTPKEPAPTTKEPAP 392

Qy        854 DTGKAEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGSGRATTAKSSGAPQ 913
               | |   |.       | : ||     | |  | :        ||     : |    ||
Db        393 TTPKEPAPT-------APKKPAPTTPKEPAPTTPKEPAPTTTKE----PSPTTPKEPAPT 441

Qy        914 DSDSSATCSADEVDEAEGGDKNRLLSPRPSLLTPTGDPRANASPQKPLDLKQLKQRAAAI 973
               : |: | : :                      ||    |  :|::|
Db        442 TTKSAPTTTKEP-------------------APTTTKSAPTTPKEP------------- 468

Qy        974 PPIQVTKVHEP----PREDA--APTKPAPPAPPPPQNLQPESDAPQQPGSSPRGKSRSPA 1027
                    |   ||    |:| |   | ||||  |  |    |:  ||           : ||
Db        469 ---SPTTTKEPAPTTPKEPAPTTPKKPAPTTPKEPAPTTPKEPAPTTTKKPAPTAPKEPA 525

Qy       1028 PPADKEAFAAEAQKLPGDPP----CWTSGLPFPVPPREVIKASPHAPDPSAFSYAPPGHP 1083
              |   ||       :||    |       |   | |  | |:   :|  | |:      |  |
Db        526 PTTPKETAPTTPKKLTPTTPEKLAPTTPEKPAPTTPEELAPTTPEEPTPTT-----PEEP 580

Qy       1084 LPLGLHDTARPVLPRPPTISNP-PPLISSAKHPSVLERQIGAISQGMSVQLHVPYSEHAK 1142
              |      | |  |: |  : |  |  :: | |:                    | :
Db        581 APT-TPKAAAPNTPKEPAPTTPKEPAPTTPKEPA------------------PTTPKET 620

Qy       1143 AP-----VGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPESLG---VPTAQEASV 1194
              ||       | |:  | |  ||| ||      ::|:|      |  :      ||
Db        621 APTTPKGTAPTTLKEPAPTTPKKPAP------KELAPTTTKEPTSTTSDKPAPTTP---- 670

Qy       1195 LRGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHGTPADVLYKGTITRIIGEDSPSR 1254
               :|||             |:|  | : | |     |   ||    |||    : | :|:

Db        671 -KGTA------------PTT--PKEPAPT-----TPKEPAPTTPKGTAPTTLKEPAPTT 709

Qy       1255 LDRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPHETA--APKRTYD 1312
                :      |               : :|  | |       : :   | |||   ||
Db        710 PKKPAPKELAP-----------TTTKGPTSTT-----SDKPAPTTPKETAPTTPKEPAP 752

Qy       1313 MMEGRVGRAISSASIEGLMGRAIPP---ERHSPHHLKEQHHIRGSITQGIPRSYVEAQED 1369
                 :                    ||    |  :|   ||    |  |  :  |
Db        753 TTPKKPAPTTPETP---------PPTTSEVSTPTTTKEPTTIHKSPDESTPE-------- 795

Qy       1370 YLRREAKLLKREGTP-----PPPPPSRDLTEAYKTQALGPLKLKPAHEGLVATVKEAGRS 1424
                       | | ||      |  |    |:        |   ||    :  | |:
Db        796 --------LSAEPTPKALENSPKEPGVPTTKT-------PAATKPE---MTTTAKD---- 833

Qy       1425 IHEIPREELRHTPE-LPLAPRPLKEGSIT----------------------QGTPLKYDT 1461
                :  :|| |||     ||:  || : |                        || |  |
Db        834 --KTTERDLRTTPETTTAAPKMTKETATTTEKTTESKITATTTQVTSTTTQDTTPFKITT 891

Qy       1462 GAS-------TTGSKKHDVRSLIGSPGRTFPPVHPLDVMADARALERACYEESLKSRPGT 1514
                :       ||  |      ::  | |              |:  :||   ::   :|
Db        892 LKTTTLAPKVTTTKKTITTTEIMNKPEET-----------AKPKDRATNSKATTPKPQK 939

Qy       1515 ASSSGGS-IARGAPVIVPELGKPRQSPLTYEDHGAPFAGHLPRGSPVTMRE--PTPRLQE 1571
                : :    :   |  :| : ||: :|                ||    || |  || |: |
Db        940 PTKAPKKPTSTKKPKTMPRVRKPKTTP------------TPRKMTSTMPELNPTSRIAE 986

Qy       1572 GSLSSS---KASQDRKLT-------------STPREIAKSPHSTVPEHHPH----PISP 1610
                | ::     : : ||              ||  : : ||  :||  |     |  |
Db        987 AMLQTTTRPNQTPNSKLVEVNPKSEDAGGAEGETPHMLLR-PHVFMPEVTPDMDYLPRVP 1045

Qy       1611 YEHLLRGVSGVDLYRSHIPLAFDPTSIPRGIPLD 1644
                : ::             |:  | |:|  | |:|
Db       1046 NQGIIIN-----------PMLSDETNICNGKPVD 1068


RESULT 33
US-10-124-557-60
; Sequence 60, Application US/10124557
; Publication No. US20020137894A1
;   GENERAL INFORMATION:
;        APPLICANT: Turner, Katherine
;                   Clark, Stephen C.
;                   Jacobs, Kenneth
;                   Hewick, Rodney M.
;                   Gesner, Thomas G.
;        TITLE OF INVENTION: Megakaryocyte Stimulating Factors
;        NUMBER OF SEQUENCES: 143
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Genetics Institute, Inc.
;             STREET: 87 CambridgePark Drive
;             CITY: Cambridge
;             STATE: Massachusetts
;             COUNTRY: U.S.A.
;             ZIP: 02140
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/124,557
;             FILING DATE: 16-Apr-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US 07/643,502
;             FILING DATE: 18-JAN-1991
;             APPLICATION NUMBER: US 07/546,114
;             FILING DATE: 29-JUN-1990
;             APPLICATION NUMBER: US 07/457,196
;             FILING DATE: 29-DEC-1989
;             APPLICATION NUMBER: US 07/390,901
;             FILING DATE: 08-AUG-1989
;        ATTORNEY/AGENT INFORMATION:
```

```
;            NAME: Cserr, Luann
;            REGISTRATION NUMBER: 31,822
;            REFERENCE/DOCKET NUMBER: GI 5190
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (617)876-1170
;            TELEFAX: (617)876-5851
;   INFORMATION FOR SEQ ID NO: 60:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 1320 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 60:
US-10-124-557-60

  Query Match             3.2%;  Score 425.5;  DB 13;  Length 1320;
  Best Local Similarity   21.2%;  Pred. No. 3.1e-11;
  Matches  261;  Conservative  127;  Mismatches  491;  Indels  355;  Gaps   53;

Qy        508 QQQPMPRSSQEEKDEKEKEKEAEKEEEKPEVENDKEDLLKEK-------TDDT-SGEDND 559
              :: | | |          |:| |:  : ||  ||:::|::  |:|         |:  || ||
Db         93 KRSPKP-------PNKKKTKKVIESEEITEVKDNKKNRTKKKPTPKPPVVDEAGSGLDNG 145

Qy        560 EKEAVASKGRKTANSQGRRKGRITRSMANEANSEEAITP-QQSAELASMELNESSRWTEE 618
              : :        | :::     :||    |    |    :: |   :::  |: :|: :    :
Db        146 DFKVTTPDTSTTQHNKVSTSPKIT--TAKPINPRPSLPPNSDTSKETSLTVNKETTVETK 203

Qy        619 EMETAKKGLLEHGRNWSAIARMVGSKTVSQCKNFYFNYKKRQNLDEILQQHKLKME---K 675
              |  |  |     |:  :  |:   |   :  |:           ::| :   | |   |
Db        204 ETTTTNKQTSTDGKEKTTSAKETQSIEKTSAKDL-------APTSKVLAKPTPKAETTTK 256

Qy        676 ERNARRKKKKAPAAASEEAAFPPVVEDEEMEASGVSGNEEEMVEEAEALHASGNEVPRGE 735
                    |:  |    | |:  |                                : |
Db        257 GPALTTPKEPTPTTPKEPASTTP------------------------------KEPT-- 283

Qy        736 CSGPATVNNSSDTESIPSPHTEAAKDTGQNGPKPPATLGADGPPPGPPTPPRRTSRAPIE 795
                | |: ::  |   |:| |  :  |    | |  |       | | ||        |
Db        284 ---PTTIKSAPTTPKEPAPTTTKSAPTTPKEPAPTTT-----KEPAPTTPKEPAPTTTKE 335

Qy        796 PTPASEATGAPTPP--PAPPSPSAPPPVVPKEEKEEETAAAPPVEEGEEQKPPAAEELAV 853
              | | :    ||| |  ||| :|  | |  |||      |     |     ::  |  :| |
Db        336 PAPTT-TKSAPTTPKEPAPTTPKKPAPTTPKE--PAPTTPKEPTPTTPKEPAPTTKEPAP 392

Qy        854 DTGKAEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGSGRATTAKSSGAPQ 913
              | |   |        | : ||     | |  | :       ||     : |     ||
Db        393 TTPKEPAPT-------APKKPAPTTPKEPAPTTPKEPAPTTTKE----PSPTTPKEPAPT 441

Qy        914 DSDSSATCSADEVDEAEGGDKNRLLSPRPSLLTPTGDPRANASPQKPLDLKQLKQRAAAI 973
              : |: | : :                  ||    |  :|::|
Db        442 TTKSAPTTTKEP-------------------APTTTKSAPTTPKEP------------- 468

Qy        974 PPIQVTKVHEP----PREDA--APTKPAPPAPPPPQNLQPESDAPQQPGSSPRGKSRSPA 1027
                    |   ||    |:| |   | ||||  |  |    |:  ||           : ||
Db        469 ---SPTTTKEPAPTTPKEPAPTTPKKPAPTTPKEPAPTTPKEPAPTTTKKPAPTAPKEPA 525

Qy       1028 PPADKEAFAAEAQKLPGDPP----CWTSGLPFPVPPREVIKASPHAPDPSAFSYAPPGHP 1083
              |   ||      :||    |       |  | |  | |:   :|  | |:      |  |
Db        526 PTTPKETAPTTPKKLTPTTPEKLAPTTPEKPAPTTPEELAPTTPEEPTPTT-----PEEP 580

Qy       1084 LPLGLHDTARPVLPRPPTISNP-PPLISSAKHPSVLERQIGAISQGMSVQLHVPYSEHAK 1142
              |      | |  |: |  : |  |  :: | |:                   | :
Db        581 APT-TPKAAAPNTPKEPAPTTPKEPAPTTPKEPA------------------PTTPKET 620

Qy       1143 AP-----VGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPESLG---VPTAQEASV 1194
              ||       | |:  | |  ||| ||      ::|:|     |  :      ||
Db        621 APTTPKGTAPTTLKEPAPTTPKKPAP------KELAPTTTKEPTSTTSDKPAPTTP---- 670

Qy       1195 LRGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHGTPADVLYKGTITRIIGEDSPSR 1254
              :|||             |:|  | : | |     |  ||    |||    : | :|:
Db        671 -KGTA------------PTT--PKEPAPT-----TPKEPAPTTPKGTAPTTLKEPAPTT 709

Qy       1255 LDRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPHETA--APKRTYD 1312
                :     |               : :|  | |       : :   | |||   ||
```

```
Db        710 PKKPAPKELAP-----------TTTKGPTSTT-----SDKPAPTTPKETAPTTPKEPAP 752

Qy       1313 MMEGRVGRAISSASIEGLMGRAIPP---ERHSPHHLKEQHHIRGSITQGIPRSYVEAQED 1369
                :                        ||   |  :|   ||   |  |  :  |
Db        753 TTPKKPAPTTPETP---------PPTTSEVSTPTTTKEPTTIHKSPDESTPE-------- 795

Qy       1370 YLRREAKLLKREGTP-----PPPPPSRDLTEAYKTQALGPLKLKPAHEGLVATVKEAGRS 1424
                  | | ||      |  |  |:        |   ||    :  | |:
Db        796 --------LSAEPTPKALENSPKEPGVPTTKT-------PAATKPE---MTTTAKD---- 833

Qy       1425 IHEIPREELRHTPE-LPLAPRPLKEGSIT----------------------QGTPLKYDT 1461
                :    :|| |||     ||:  || : |                      || | |
Db        834 --KTTERDLRTTPETTTAAPKMTKETATTTEKTTESKITATTTQVTSTTTQDTTPFKITT 891

Qy       1462 GAS-------TTGSKKHDVRSLIGSPGRTFPPVHPLDVMADARALERACYEESLKSRPGT 1514
                :       ||  |       ::  |  |               |:  :||   ::   :|
Db        892 LKTTTLAPKVTTTKKTITTTEIMNKPEET-----------AKPKDRATNSKATTPKPQK 939

Qy       1515 ASSSGGS-IARGAPVIVPELGKPRQSPLTYEDHGAPFAGHLPRGSPVTMRE--PTPRLQE 1571
                : :      :   |  :| : ||: :|              ||    || |  || |: |
Db        940 PTKAPKKPTSTKKPKTMPRVRKPKTTP------------TPRKMTSTMPELNPTSRIAE 986

Qy       1572 GSLSSS---KASQDRKLT-------------STPREIAKSPHSTVPEHHPH----PISP 1610
                | ::     : : ||               ||  : : ||  :||  |      | |
Db        987 AMLQTTTRPNQTPNSKLVEVNPKSEDAGGAEGETPHMLLR-PHVFMPEVTPDMDYLPRVP 1045

Qy       1611 YEHLLRGVSGVDLYRSHIPLAFDPTSIPRGIPLD 1644
                : ::              |:  | |:|  | |:|
Db       1046 NQGIIIN-----------PMLSDETNICNGKPVD 1068


RESULT 34
US-10-124-557-40
; Sequence 40, Application US/10124557
; Publication No. US20020137894A1
;   GENERAL INFORMATION:
;        APPLICANT: Turner, Katherine
;                   Clark, Stephen C.
;                   Jacobs, Kenneth
;                   Hewick, Rodney M.
;                   Gesner, Thomas G.
;        TITLE OF INVENTION: Megakaryocyte Stimulating Factors
;        NUMBER OF SEQUENCES: 143
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Genetics Institute, Inc.
;             STREET: 87 CambridgePark Drive
;             CITY: Cambridge
;             STATE: Massachusetts
;             COUNTRY: U.S.A.
;             ZIP: 02140
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/124,557
;             FILING DATE: 16-Apr-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US 07/643,502
;             FILING DATE: 18-JAN-1991
;             APPLICATION NUMBER: US 07/546,114
;             FILING DATE: 29-JUN-1990
;             APPLICATION NUMBER: US 07/457,196
;             FILING DATE: 29-DEC-1989
;             APPLICATION NUMBER: US 07/390,901
;             FILING DATE: 08-AUG-1989
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Cserr, Luann
;             REGISTRATION NUMBER: 31,822
;             REFERENCE/DOCKET NUMBER: GI 5190
;        TELECOMMUNICATION INFORMATION:
```

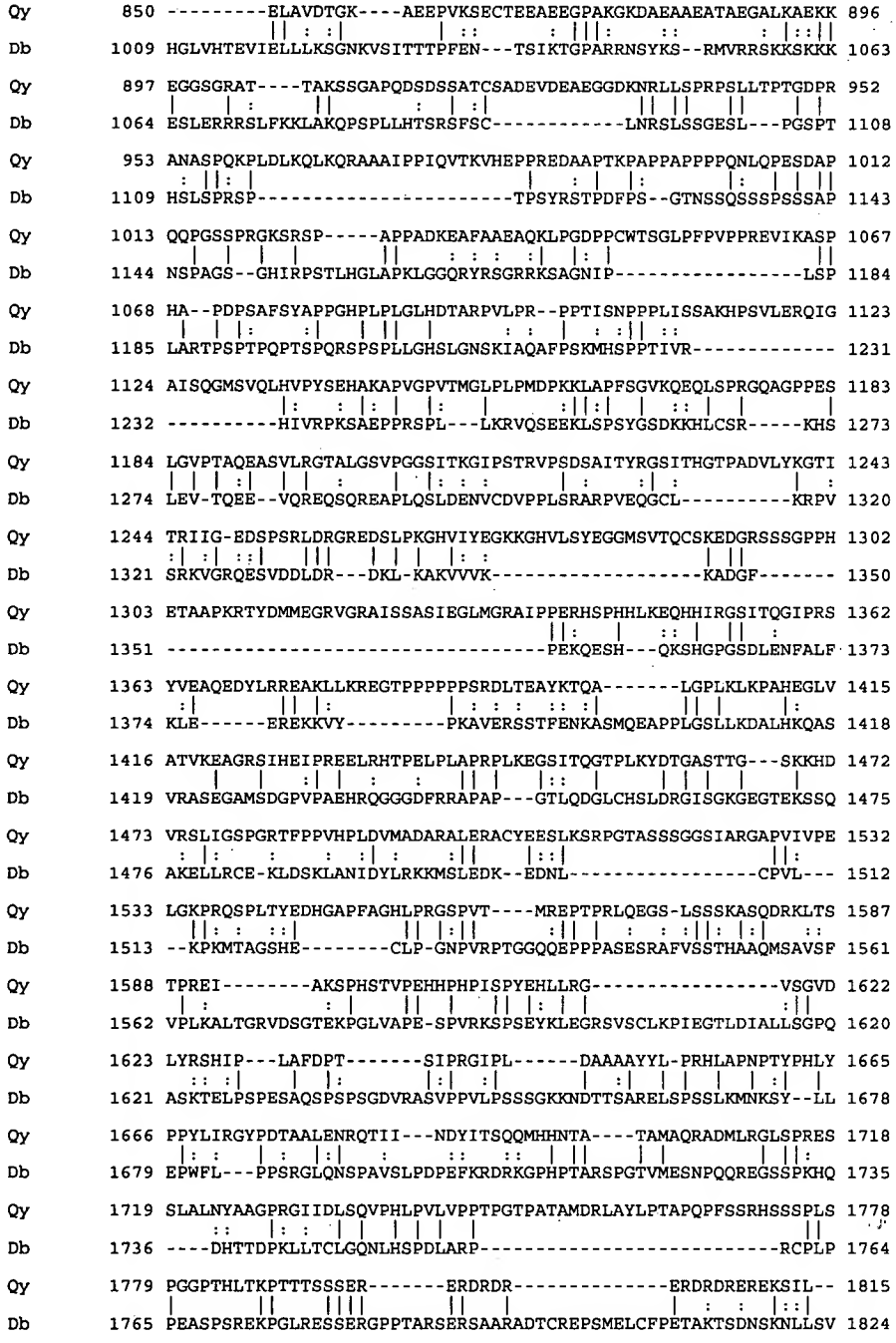
```
;            TELEPHONE: (617)876-1170
;            TELEFAX: (617)876-5851
;   INFORMATION FOR SEQ ID NO: 40:
;        SEQUENCE CHARACTERISTICS:
;            LENGTH: 1361 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 40:
US-10-124-557-40

  Query Match             3.2%;  Score 425.5;  DB 13;  Length 1361;
  Best Local Similarity   21.2%;  Pred. No. 3.2e-11;
  Matches  261;  Conservative  127;  Mismatches  491;  Indels  355;  Gaps   53;

Qy        508 QQQPMPRSSQEEKDEKEKEKEAEKEEEKPEVENDKEDLLKEK-------TDDT-SGEDND 559
              :: | .|         |:| |:  : ||  ||:::|::  |:|         |:  || ||
Db        134 KRSPKP-------PNKKKTKKVIESEEITEVKDNKKNRTKKKPTPKPPVVDEAGSGLDNG 186

Qy        560 EKEAVASKGRKTANSQGRRKGRITRSMANEANSEEAITP-QQSAELASMELNESSRWTEE 618
              : :        | :::     :||    |    |   :: |   :::   |: :|: :    :
Db        187 DFKVTTPDTSTTQHNKVSTSPKIT--TAKPINPRPSLPPNSDTSKETSLTVNKETTVETK 244

Qy        619 EMETAKKGLLEHGRNWSAIARMVGSKTVSQCKNFYFNYKKRQNLDEILQQHKLKME---K 675
              |  |  |     |:  :  |:   |   :  |:          ::| :   | |   |
Db        245 ETTTTNKQTSTDGKEKTTSAKETQSIEKTSAKDL-------APTSKVLAKPTPKAETTTK 297

Qy        676 ERNARRKKKKAPAAASEEAAFPPVVEDEEMEASGVSGNEEEMVEEAEALHASGNEVPRGE 735
                    |:  |    | |:  |                                 : |
Db        298 GPALTTPKEPTPTTPKEPASTTP------------------------------KEPT-- 324

Qy        736 CSGPATVNNSSDTESIPSPHTEAAKDTGQNGPKPPATLGADGPPPGPPTPPRRTSRAPIE 795
                | |: ::  |   |:| |  :  |   | | |       | | ||          |
Db        325 ---PTTIKSAPTTPKEPAPTTTKSAPTTPKEPAPTTT-----KEPAPTTPKEPAPTTTKE 376

Qy        796 PTPASEATGAPTPP--PAPPSPSAPPPVVPKEEKEEETAAAPPVEEGEEQKPPAAEELAV 853
              | | :    ||| |  ||| :|  | |  |||     |    |     ::  |  :| |
Db        377 PAPTT-TKSAPTTPKEPAPTTPKKPAPTTPKE--PAPTTPKEPTPTTPKEPAPTTKEPAP 433

Qy        854 DTGKAEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGSGRATTAKSSGAPQ 913
              | |  |        | : ||      | |  | :       ||    : |     ||
Db        434 TTPKEPAPT-------APKKPAPTTPKEPAPTTPKEPAPTTTKE----PSPTTPKEPAPT 482

Qy        914 DSDSSATCSADEVDEAEGGDKNRLLSPRPSLLTPTGDPRANASPQKPLDLKQLKQRAAAI 973
              : |: | : :                 ||     | :|::|
Db        483 TTKSAPTTTKEP-------------------APTTTKSAPTTPKEP------------- 509

Qy        974 PPIQVTKVHEP----PREDA--APTKPAPPAPPPPQNLQPESDAPQQPGSSPRGKSRSPA 1027
                    |   ||    |:| |   | ||||  |  |    |:  ||          : ||
Db        510 ---SPTTTKEPAPTTPKEPAPTTPKKPAPTTPKEPAPTTPKEPAPTTTKKPAPTAPKEPA 566

Qy       1028 PPADKEAFAAEAQKLPGDPP----CWTSGLPFPVPPREVIKASPHAPDPSAFSYAPPGHP 1083
              |   ||       :||     |       |   | |  | |:   :|  | |:       |  |
Db        567 PTTPKETAPTTPKKLTPTTPEKLAPTTPEKPAPTTPEELAPTTPEEPTPTT-----PEEP 621

Qy       1084 LPLGLHDTARPVLPRPPTISNP-PPLISSAKHPSVLERQIGAISQGMSVQLHVPYSEHAK 1142
              |       | |  |: |  : |  |  :: | |:                   | :
Db        622 APT-TPKAAAPNTPKEPAPTTPKEPAPTTPKEPA------------------PTTPKET 661

Qy       1143 AP-----VGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPESLG---VPTAQEASV 1194
              ||       | |:  | |  ||| ||      ::|:|    |  :      ||
Db        662 APTTPKGTAPTTLKEPAPTTPKKPAP------KELAPTTTKEPTSTTSDKPAPTTP---- 711

Qy       1195 LRGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHGTPADVLYKGTITRIIGEDSPSR 1254
               :|||             |:|  | : | |      |   ||    |||    : | :|:
Db        712 -KGTA------------PTT--PKEPAPT-----TPKEPAPTTPKGTAPTTLKEPAPTT 750

Qy       1255 LDRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPHETA--APKRTYD 1312
                :     |                 : :|  | |       : :   | |||   ||
Db        751 PKKPAPKELAP------------TTTKGPTSTT-----SDKPAPTTPKETAPTTPKEPAP 793

Qy       1313 MMEGRVGRAISSASIEGLMGRAIPP---ERHSPHHLKEQHHIRGSITQGIPRSYVEAQED 1369
                  :                   ||   | :|   ||   |  |  :  |

Db        794 TTPKKPAPTTPETP---------PPTTSEVSTPTTTKEPTTIHKSPDESTPE-------- 836

Qy       1370 YLRREAKLLKREGTP-----PPPPPSRDLTEAYKTQALGPLKLKPAHEGLVATVKEAGRS 1424
                       | | ||      |  |    |:       |  ||    :  | |:
Db        837 --------LSAEPTPKALENSPKEPGVPTTKT-------PAATKPE---MTTTAKD---- 874

Qy       1425 IHEIPREELRHTPE-LPLAPRPLKEGSIT----------------------QGTPLKYDT 1461
                :    :|| |||    ||:  || : |                        || | |
Db        875 --KTTERDLRTTPETTTAAPKMTKETATTTEKTTESKITATTTQVTSTTTQDTTPFKITT 932

Qy       1462 GAS-------TTGSKKHDVRSLIGSPGRTFPPVHPLDVMADARALERACYEESLKSRPGT 1514
                :       ||  |      ::  |  |             |:  :||   ::   :|
Db        933 LKTTTLAPKVTTTKKTITTTEIMNKPEET------------AKPKDRATNSKATTPKPQK 980

Qy       1515 ASSSGGS-IARGAPVIVPELGKPRQSPLTYEDHGAPFAGHLPRGSPVTMRE--PTPRLQE 1571
                : :     :   |  :| : ||: :|             ||    || |  || |: |
Db        981 PTKAPKKPTSTKKPKTMPRVRKPKTTP-------------TPRKMTSTMPELNPTSRIAE 1027

Qy       1572 GSLSSS---KASQDRKLT-------------STPREIAKSPHSTVPEHHPH----PISP 1610
                | ::     : : ||              ||  : : ||  :||  |      |  |
Db       1028 AMLQTTTRPNQTPNSKLVEVNPKSEDAGGAEGETPHMLLR-PHVFMPEVTPDMDYLPRVP 1086

Qy       1611 YEHLLRGVSGVDLYRSHIPLAFDPTSIPRGIPLD 1644
                : ::              |:  | |:|  | |:|
Db       1087 NQGIIIN-----------PMLSDETNICNGKPVD 1109


RESULT 35
US-10-042-865-12
; Sequence 12, Application US/10042865
; Publication No. US20040029216A1
; GENERAL INFORMATION:
;  APPLICANT: Padigaru, Muralidhara
;  APPLICANT:  Li, Li
;  APPLICANT:  Zerhusen, Bryan D
;  APPLICANT:  Casman, Stacie J
;  APPLICANT:  Shenoy, Suresh G
;  APPLICANT:  Spytek, Kimberly
;  APPLICANT:  Zhong, Mei
;  APPLICANT:  Gangolli, Esha A
;  APPLICANT:  Burgess, Catherine E
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Vernet, Corine A.M
;  APPLICANT:  Taylor, Sarah
;  APPLICANT:  Tchernev, Velizar T
;  APPLICANT:  Miller, Charles E
;  APPLICANT:  Guo, Xiaojia
;  APPLICANT:  Boldog, Ference L
;  APPLICANT:  Grosse, William M
;  APPLICANT:  Alsobrook II, John P
;  APPLICANT:  Gerlach, Valerie L
;  APPLICANT:  Edinger, Shlomit R
;  APPLICANT:  Rothenberg, Mark E
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  MacDougall, John
;  APPLICANT:  Malyankar, Uriel M
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  Peyman, John
;  APPLICANT:  Smithson, Glennda
;  APPLICANT:  Gunther, Erik
;  APPLICANT:  Stone, David
;  TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
;  TITLE OF INVENTION:  Using the Same
;  FILE REFERENCE: 21402-537
;  CURRENT APPLICATION NUMBER: US/10/042,865
;  CURRENT FILING DATE:  2002-05-17
;  PRIOR APPLICATION NUMBER: 60/260,417
;  PRIOR FILING DATE: 2001-01-09
;  PRIOR APPLICATION NUMBER: 60/260,831
;  PRIOR FILING DATE: 2001-01-10
;  PRIOR APPLICATION NUMBER: 60/272,338
;  PRIOR FILING DATE: 2001-02-28
;  PRIOR APPLICATION NUMBER: 60/274,876
```

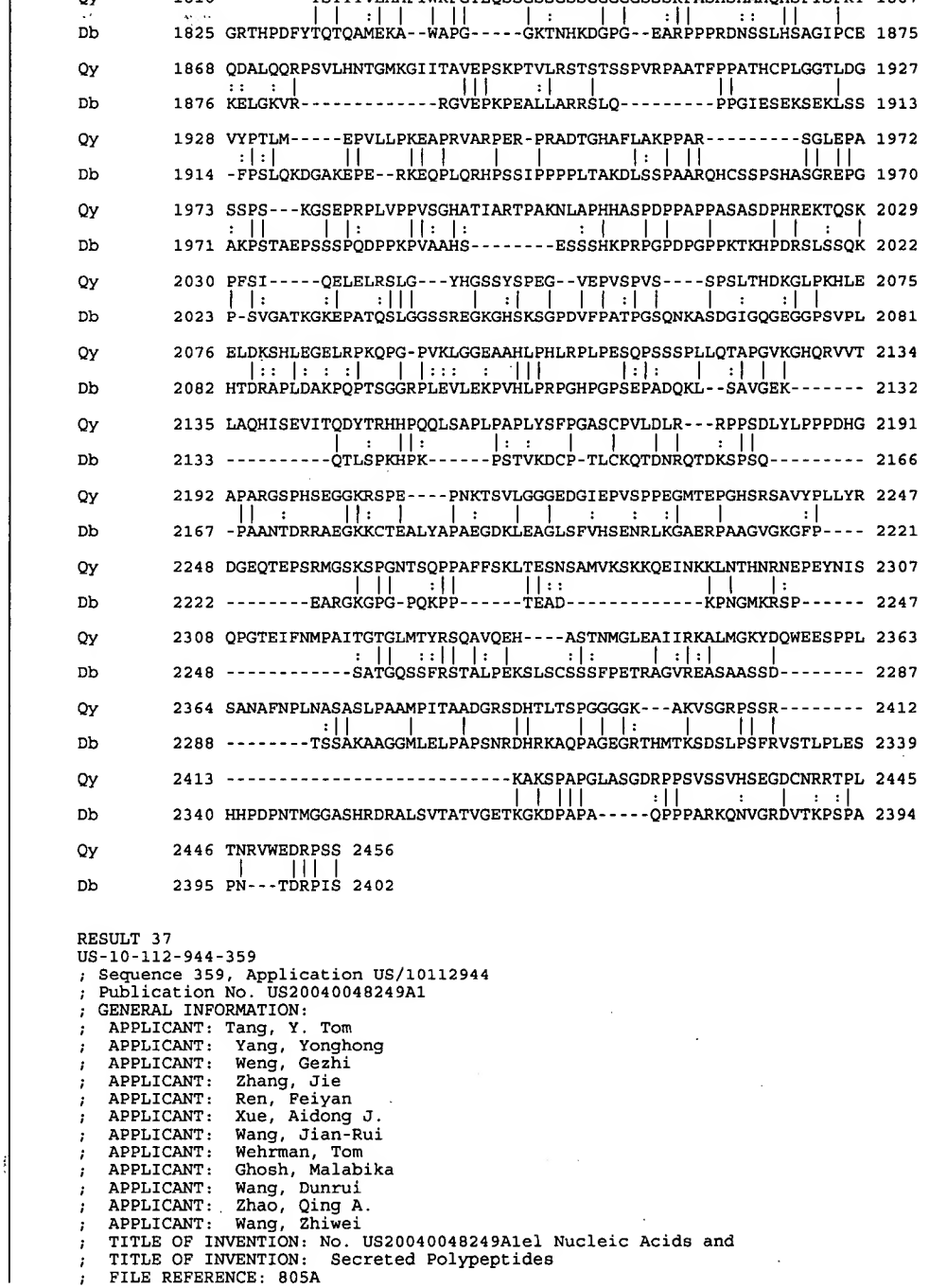
```
;  PRIOR FILING DATE: 2001-03-09
;  PRIOR APPLICATION NUMBER: 60/284,704
;  PRIOR FILING DATE: 2001-04-18
;  NUMBER OF SEQ ID NOS: 264
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
;   LENGTH: 2545
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-042-865-12

  Query Match             3.2%;  Score 425.5;  DB 12;  Length 2545;
  Best Local Similarity   19.8%;  Pred. No. 6.7e-11;
  Matches  549;  Conservative  308;  Mismatches  979;  Indels  935;  Gaps  139;

Qy         38 GLLEYQHH-----SRDYASHLSPGSI-----IQPQRRRPSLLSEFQPGNERSQELHLRPE 87
Db        331 GVLSFTHHQIIELARDCLDKSHQGLITSRYFLELQHKLDKLL----------QEAHDRSE 380

Qy         88 SHSYLPELGKSEMEFIESKRPRLELLPDPLLRPSPLLATGQPAGSEDLTKDRSLTGKLEP 147
Db        381 S---------GELAFIKQLVRKILIV---IARPARLLEC------------------LE- 409

Qy        148 VSPPSPPHTDPE-----LELVPPRLSKEELIQNMDRVDREITMVEQQISKLKKKQQQLEE 202
Db        410 --------FDPEEFYYLLEAAEGHAKEGQGIKT--DIPRYI------ISQLGLNKDPLEE 453

Qy        203 EA---------AKPPEPEKPVSPPPIESKHRSLVQIIYDENRKKAEAAHRILEGLGPQVE 253
Db        454 MAHLGNYDSGTAETPETDESVSSSNASLKLR----------RKPRE-------------- 489

Qy        254 LPLYNQPSDTRQYHENIKI--NQAMRKKLILYFKRRNHARKQWKQKFCQRY-----DQLM 306
Db        490 -------SD----FETIKLISNGAYG---AVYFVRHKESRQRFAMKKINKQNLILRNQIQ 535

Qy        307 EA-LEKKVERIENNPRRRAKESKVREYYEKQFPEIRKQRELQERMQSRVG------QRGS 359
Db        536 QAFVERDILTFAENP-----------FVVSMYCSFETRRHLCMVMEYVEGGDCATLMKNM 584

Qy        360 G-LSMSAARSEHEVSEIIDGLSEQENLEKQMRQLAVIPPMLYDADQQRIKFINMN----G 414
Db        585 GPLPVDMARM--YFAETVLALEYLHNYGIVHRDLK--PDSLLVTSMGHIKLTDFGLSKVG 640

Qy        415 LMADPMKVYK-----------DRQVMNMWSEQEKETFREKFMQHPKNFGLIASFLERKTV 463
Db        641 LMSMTTNLYEGHIEKDAREFLDKQVCGTPEYIAPEVILRQGYGKPVDWWAMGIILYEFLV 700

Qy        464 AECVLYYYLT----------------KKNE----NYKSLV-----RRSYRRRGKSQQQQQ 498
Db        701 G-CVPFFGDTPEELFGQVISDEINWPEKDEAPPPDAQDLITLLLRQNPLERLGTGGAYEV 759

Qy        499 QQQQQQQQQQQQPMPRSSQEEKDEKEKEKEAEKEEEKPEVENDKEDLLKEKTDDTSGED- 557
Db        760 KQHRFFRSLDWNSLLRQKAEFIPQLESEDDTSYFDTRSEKYHHME---TEEEDDTNDEDF 816

Qy        558 NDEKEAVAS---KGRKTANSQGRRKGRITRSMANEANSEEAITPQQSAELASME-LNESS 613
Db        817 NVEIRQFSSCSHRFSKVFSS----IDRITQNSAEE--KEDSVDKTKSTTLPSTETLSWSS 870

Qy        614 RWTE-EEMETAKKGLLEHGRNWSAIARMVGSKTVSQCKNFYFNYKKRQNLDEILQQHKLK 672
Db        871 EYSEMQQLSTSNSSDTESNR----------------------HKLSSGL---LPKLAIS 904

Qy        673 MEKERNARRKKKAPAAASEEAAFPPVVEDEEMEASGVSGNEEEMVEEAEALHASGNEV- 731
Db        905 TEGEQD---EAASCPGDPHEEPGKPALPPEE------CAQEEPEVTTPASTISSSTLSVG 955

Qy        732 ----------PRGECSGPATVNNSSDTESIPSPH--------TEAAKDTGQNGPKPPATL 773
Db        956 SFSEHLDQINGRSECVD--STDNSSKPSSEPASHMARQRLESTEKKKISG----KVTKSL 1009

Qy        774 GADG-----------PPPGPPTPPRRTSRAPIEPTPASEATGAPTPPPAPPSPSAPPPV 821
Db       1010 SASALSLMIPGDMFAVSPLGSPMSPHSLS-----SDPSSSRDSSPSRDSSAASASPHQPI 1064

Qy        822 VPKEEKEEE--TAAAPPVEEGE-------------EQKPPAAE---------------- 849
Db       1065 VIHSSGKNYGFTIRAIRVYVGDSDIYTVHHIVWNVEEGSPACQAGLKAGDLITHINGEPV 1124

Qy        850 ---------ELAVDTGK----AEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEKK 896
Db       1125 HGLVHTEVIELLLKSGNKVSITTTPFEN---TSIKTGPARRNSYKS--RMVRRSKKSKKK 1179

Qy        897 EGGSGRAT----TAKSSGAPQDSDSSATCSADEVDEAEGGDKNRLLSPRPSLLTPTGDPR 952
Db       1180 ESLERRRSLFKKLAKQPSPLLHTSRSFSC------------LNRSLSSGESL---PGSPT 1224

Qy        953 ANASPQKPLDLKQLKQRAAAIPPIQVTKVHEPPREDAAPTKPAPPAPPPPQNLQPESDAP 1012
Db       1225 HSLSPRSP------------------------TPSYRSTPDFPS--GTNSSQSSSPSSSAP 1259

Qy       1013 QQPGSSPRGKSRSP-----APPADKEAFAAEAQKLPGDPPCWTSGLPFPVPPREVIKASP 1067
Db       1260 NSPAGS--GHIRPSTLHGLAPKLGGQRYRSGRRKSAGNIP-----------------LSP 1300

Qy       1068 HA--PDPSAFSYAPPGHPLPLGLHDTARPVLPR--PPTISNPPPLISSAKHPSVLERQIG 1123
Db       1301 LARTPSPTPQPTSPQRSPSPLLGHSLGNSKIAQAFPSKMHSPPTIVR------------ 1347

Qy       1124 AISQGMSVQLHVPYSEHAKAPVGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPES 1183
Db       1348 ----------HIVRPKSAEPPRSPL---LKRVQSEEKLSPSYGSDKKHLCSR-----KHS 1389

Qy       1184 LGVPTAQEASVLRGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHGTPADVLYKGTI 1243
Db       1390 LEV-TQEE--VQREQSQREAPLQSLDENVCDVPPLSRARPVEQGCL----------KRPV 1436

Qy       1244 TRIIG-EDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPH 1302
Db       1437 SRKVGRQESVDDLDR---DKL-KAKVVVK------------------KADGF------- 1466

Qy       1303 ETAAPKRTYDMMEGRVGRAISSASIEGLMGRAIPPERHSPHHLKEQHHIRGSITQGIPRS 1362
Db       1467 ----------------------------------PEKQESH---QKSHGPGSDLENFALF 1489

Qy       1363 YVEAQEDYLRREAKLLKREGTPPPPPPSRDLTEAYKTQA-------LGPLKLKPAHEGLV 1415
Db       1490 KLE------EREKKVY----------PKAVERSSTFENKASMQEAPPLGSLLKDALHKQAS 1534

Qy       1416 ATVKEAGRSIHEIPREELRHTPELPLAPRPLKEGSITQGTPLKYDTGASTTG---SKKHD 1472
Db       1535 VRASEGAMSDGPVPAEHRQGGGDFRRAPAP---GTLQDGLCHSLDRGISGKGEGTEKSSQ 1591

Qy       1473 VRSLIGSPGRTFPPVHPLDVMADARALERACYEESLKSRPGTASSSGGSIARGAPVIVPE 1532
Db       1592 AKELLRCE-KLDSKLANIDYLRKKMSLEDK--EDNL-----------------CPVL--- 1628

Qy       1533 LGKPRQSPLTYEDHGAPFAGHLPRGSPVT----MREPTPRLQEGS-LSSSKASQDRKLTS 1587
Db       1629 --KPKMTAGSHE--------CLP-GNPVRPTGGQQEPPPASESRAFVSSTHAAQMSAVSF 1677

Qy       1588 TPREI--------AKSPHSTVPEHHPHPISPYEHLLRG-----------------VSGVD 1622
Db       1678 VPLKALTGRVDSGTEKPGLVAPE-SPVRKSPSEYKLEGRSVSCLEPIEGTLDIALLSGPQ 1736

Qy       1623 LYRSHIP---LAFDPT-------SIPRGIPL------DAAAAYYL-PRHLAPNPTYPHLY 1665
Db       1737 ASKTELPSPESAQSPSPSGDVRASVPPVLPSSSGKKNDTTSARELSPSSLKMNKSY--LL 1794

Qy       1666 PPYLIRGYPDTAALENRQTII---NDYITSQQMHHNTA----TAMAQRADMLRGLSPRES 1718
Db       1795 EPWFL---PPSRGLQNSPAVSLPDPEFKRDRKGPHPTARSPGTVMESNPQQREGSSPKHQ 1851

Qy       1719 SLALNYAAGPRGIIDLSQVPHLPVLVPPTPGTPATAMDRLAYLPTAPQPFSSRHSSSPLS 1778
Db       1852 ----DHTTDPKLLTCLGQNLHSPDLARP--------------------------RCPLP 1880
```

```
Qy       1779 PGGPTHLTKPTTTSSSER-------ERDRDR--------------ERDRDREREKSIL-- 1815
Db       1881 PEASPSREKPGLRESSERGPPTARSERSAARADTCREPSMELCFPETAKTSDNSKNLLSV 1940

Qy       1816 --------TSTTTVEHAPIWRPGTEQSSGSSGSSGGGGGSSSRPASHSHAHQHSPISPRT 1867
Db       1941 GRTHPDFYTQTQAMEKA--WAPG-----GKTNHKDGPG--EARPPPRDNSSLHSAGIPCE 1991

Qy       1868 QDALQQRPSVLHNTGMKGIITAVEPSKPTVLRSTSTSSPVRPAATFPPATHCPLGGTLDG 1927
Db       1992 KELGKVR------------RGVEPKPEALLARRSLQ---------PPGIESEKSEKLSS 2029

Qy       1928 VYPTLM-----EPVLLPKEAPRVARPER-PRADTGHAFLAKPPAR---------SGLEPA 1972
Db       2030 -FPSLQKDGAKEPE--RKEQPLQRHPSSIPPPPLTAKDLSSPAARQHCSSPSHASGREPG 2086

Qy       1973 SSPS---KGSEPRPLVPPVSGHATIARTPAKNLAPHHASPDPPAPPASASDPHREKTQSK 2029
Db       2087 AKPSTAEPSSSPQDPPKPVAAHS--------ESSSHKPRPGPDPGPPKTKHPDRSLSSQK 2138

Qy       2030 PFSI-----QELELRSLG---YHGSSYSPEG--VEPVSPVS----SPSLTHDKGLPKHLE 2075
Db       2139 P-SVGATKGKEPATQSLGGSSREGKGHSKSGPDVFPATPGSQNKASDGIGQGEGGPSVPL 2197

Qy       2076 ELDKSHLEGELRPKQPG-PVKLGGEAAHLPHLRPLPESQPSSSPLLQTAPGVKGHQRVVT 2134
Db       2198 HTDRAPLDAKPQPTSGGRPLEVLEKPVHLPRPGHPGPSEPADQKL--SAVGEK------- 2248

Qy       2135 LAQHISEVITQDYTRHHPQQLSAPLPAPLYSFPGASCPVLDLR---RPPSDLYLPPPDHG 2191
Db       2249 ----------QTLSPKHPK------PSTVKDCP-TLCKQTDNRQTDKSPSQ--------- 2282

Qy       2192 APARGSPHSEGGKRSPE----PNKTSVLGGGEDGIEPVSPPEGMTEPGHSRSAVYPLLYR 2247
Db       2283 -PAANTDRRAEGKKCTEALYAPAEGDKLEAGLSFVHSENRLKGAERPAAGVGKGFP---- 2337

Qy       2248 DGEQTEPSRMGSKSPGNTSQPPAFFSKLTESNSAMVKSKKQEINKKLNTHNRNEPEYNIS 2307
Db       2338 --------EARGKGPG-PQKPP------TEAD-------------KPNGMKRSP------ 2363

Qy       2308 QPGTEIFNMPAITGTGLMTYRSQAVQEH----ASTNMGLEAIIRKALMGKYDQWEESPPL 2363
Db       2364 ------------SATGQSSFRSTALPEKSLSCSSSFPETRAGVREASAASSD-------- 2403

Qy       2364 SANAFNPLNASASLPAAMPITAADGRSDHTLTSPGGGGK---AKVSGRPSSR-------- 2412
Db       2404 --------TSSAKAAGGMLELPAPSNRDHRKAQPAGEGRTHMTKSDSLPSFRVSTLPLES 2455

Qy       2413 ----------------------------KAKSPAPGLASGDRPPSVSSVHSEGDCNRRTPL 2445
Db       2456 HHPDPNTMGGASHRDRALSVTATVGETKGKDPAPA-----QPPPARKQNVGRDVTKPSPA 2510

Qy       2446 TNRVWEDRPSS 2456
Db       2511 PN---TDRPIS 2518
```

```
RESULT 36
US-10-377-035-17
; Sequence 17, Application US/10377035
; Publication No. US20030224406A1
; GENERAL INFORMATION:
;  APPLICANT: Exelixis, Inc.
;  TITLE OF INVENTION: MBCATs AS MODIFIERS OF THE BETA-CATENIN PATHWAY AND METHODS OF
;  TITLE OF INVENTION:  USE
;  FILE REFERENCE: EX03-011C
;  CURRENT APPLICATION NUMBER: US/10/377,035
;  CURRENT FILING DATE:  2003-02-28
;  PRIOR APPLICATION NUMBER: 60/361,242
;  PRIOR FILING DATE: 2002-03-01
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: PatentIn version 3.2
```

```
; SEQ ID NO 17
;   LENGTH: 2429
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-377-035-17

  Query Match             3.2%;  Score 424.5;  DB 12;  Length 2429;
  Best Local Similarity   19.8%;  Pred. No. 7e-11;
  Matches  549;  Conservative  308;  Mismatches  979;  Indels  935;  Gaps  139;

Qy         38 GLLEYQHH-----SRDYASHLSPGSI-----IQPQRRRPSLLSEFQPGNERSQELHLRPE 87
Db        215 GVLSFTHHQIIELARDCLDKSHQGLITSRYFLELQHKLDKLL----------QEAHDRSE 264

Qy         88 SHSYLPELGKSEMEFIESKRPRLELLPDPLLRPSPLLATGQPAGSEDLTKDRSLTGKLEP 147
Db        265 S---------GELAFIKQLVRKILIV---IARPARLLEC-----------------LE- 293

Qy        148 VSPPSPPHTDPE-----LELVPPRLSKEELIQNMDRVDREITMVEQQISKLKKKQQQLEE 202
Db        294 --------FDPEEFYYLLEAAEGHAKEGQGIKT--DIPRYI------ISQLGLNKDPLEE 337

Qy        203 EA---------AKPPEPEKPVSPPPIESKHRSLVQIIYDENRKKAEAAHRILEGLGPQVE 253
Db        338 MAHLGNYDSGTAETPETDESVSSSNASLKLR----------RKPRE-------------- 373

Qy        254 LPLYNQPSDTRQYHENIKI--NQAMRKKLILYFKRRNHARKQWKQKFCQRY-----DQLM 306
Db        374 -------SD----FETIKLISNGAYG---AVYFVRHKESRQRFAMKKINKQNLILRNQIQ 419

Qy        307 EA-LEKKVERIENNPRRRAKESKVREYYEKQFPEIRKQRELQERMQSRVG------QRGS 359
Db        420 QAFVERDILTFAENP----------FVVSMYCSFETRRHLCMVMEYVEGGDCATLMKNM 468

Qy        360 G-LSMSAARSEHEVSEIIDGLSEQENLEKQMRQLAVIPPMLYDADQQRIKFINMN----G 414
Db        469 GPLPVDMARM--YFAETVLALEYLHNYGIVHRDLK--PDNLLVTSMGHIKLTDFGLSKVG 524

Qy        415 LMADPMKVYK----------DRQVMNMWSEQEKETFREKFMQHPKNFGLIASFLERKTV 463
Db        525 LMSMTTNLYEGHIEKDAREFLDKQVCGTPEYIAPEVILRQGYGKPVDWWAMGIILYEFLV 584

Qy        464 AECVLYYYLT---------------KKNE----NYKSLV-----RRSYRRRGKSQQQQQ 498
Db        585 G-CVPFFGDTPEELFGQVISDEINWPEKDEAPPPDAQDLITLLLRQNPLERLGTGGAYEV 643

Qy        499 QQQQQQQQQQQQPMPRSSQEEKDEKEKEKEAEKEEEKPEVENDKEDLLKEKTDDTSGED- 557
Db        644 KQHRFFRSLDWNSLLRQKAEFIPQLESEDDTSYFDTRSEKYHHME---TEEEDDTNDEDF 700

Qy        558 NDEKEAVAS---KGRKTANSQGRRKGRITRSMANEANSEEAITPQQSAELASME-LNESS 613
Db        701 NVEIRQFSSCSHRFSKVFSS----IDRITQNSAEE--KEDSVDKTKSTTLPSTETLSWSS 754

Qy        614 RWTE-EEMETAKKGLLEHGRNWSAIARMVGSKTVSQCKNFYFNYKKRQNLDEILQQHKLK 672
Db        755 EYSEMQQLSTSNSSDTESNR----------------------HKLSSGL---LPKLAIS 788

Qy        673 MEKERNARRKKKKAPAAASEEAAFPPVVEDEEMEASGVSGNEEEMVEEAEALHASGNEV- 731
Db        789 TEGEQD---EAASCPGDPHEEPGKPALPPEE------CAQEEPEVTTPASTISSSTLSVG 839

Qy        732 ----------PRGECSGPATVNNSSDTESIPSPH--------TEAAKDTGQNGPKPPATL 773
Db        840 SFSEHLDQINGRSECVD--STDNSSKPSSEPASHMARQRLESTEKKKISG----KVTKSL 893

Qy        774 GADG------------PPPGPPTPPRRTSRAPIEPTPASEATGAPTPPPAPPSPSAPPPV 821
Db        894 SASALSLMIPGDMFAVSPLGSPMSPHSLS-----SDPSSSRDSSPSRDSSAASASPHQPI 948

Qy        822 VPKEEKEEE--TAAAPPVEEGE-------------EQKPPAAE----------------- 849
Db        949 VIHSSGKNYGFTIRAIRVYVGDSDIYTVHHIVWNVEEGSPACQAGLKAGDLITHINGEPV 1008
```



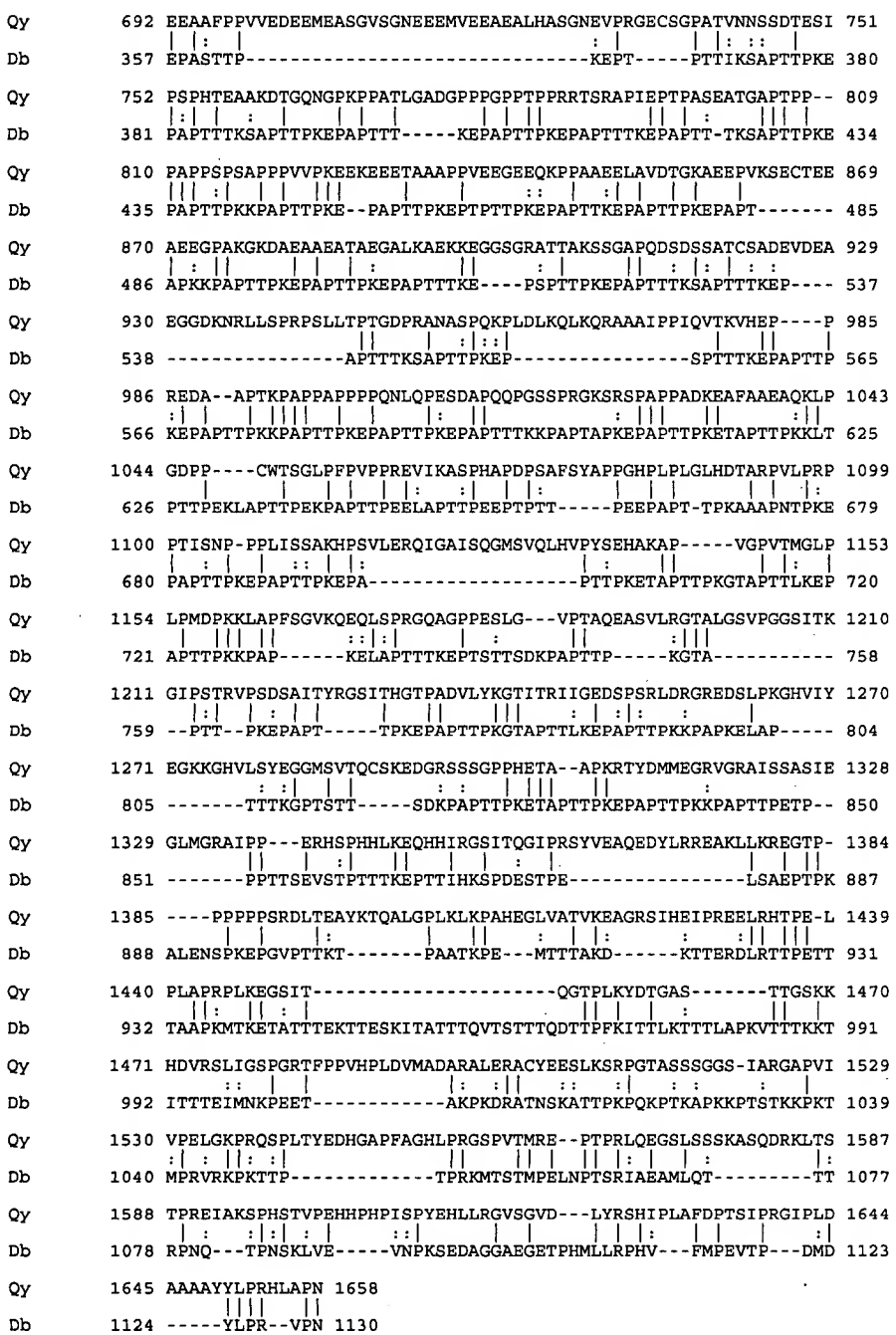


RESULT 37
US-10-112-944-359
; Sequence 359, Application US/10112944
; Publication No. US20040048249A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Yang, Yonghong
; APPLICANT: Weng, Gezhi
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Peiyan
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Malabika
; APPLICANT: Wang, Dunrui
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Zhiwei
; TITLE OF INVENTION: No. US20040048249A1el Nucleic Acids and
; TITLE OF INVENTION: Secreted Polypeptides
; FILE REFERENCE: 805A

```
;  CURRENT APPLICATION NUMBER: US/10/112,944
;  CURRENT FILING DATE:  2002-03-28
;  PRIOR APPLICATION NUMBER: US 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  PRIOR APPLICATION NUMBER: US 09/491,404
;  PRIOR FILING DATE: 2000-01-25
;  PRIOR APPLICATION NUMBER: US 09/496,914
;  PRIOR FILING DATE: 2000-02-03
;  PRIOR APPLICATION NUMBER: US 09/515,126
;  PRIOR FILING DATE: 2000-02-28
;  PRIOR APPLICATION NUMBER: US 09/519,705
;  PRIOR FILING DATE: 2000-03-07
;  PRIOR APPLICATION NUMBER: US 09/540,217
;  PRIOR FILING DATE: 2000-03-31
;  PRIOR APPLICATION NUMBER: US 09/552,929
;  PRIOR FILING DATE: 2000-04-18
;  PRIOR APPLICATION NUMBER: US 09/577,408
;  PRIOR FILING DATE: 2000-05-18
;  NUMBER OF SEQ ID NOS: 924
;  SOFTWARE: pt_FL_genes Version 5.0
; SEQ ID NO 359
;   LENGTH: 2432
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-112-944-359

  Query Match             3.2%;  Score 424.5;  DB 12;  Length 2432;
  Best Local Similarity   19.8%;  Pred. No. 7.1e-11;
  Matches  549;  Conservative  308;  Mismatches  979;  Indels  935;  Gaps  139;

Qy         38 GLLEYQHH-----SRDYASHLSPGSI-----IQPQRRRPSLLSEFQPGNERSQELHLRPE 87
              |:| : ||     :||       | |     :: | :   ||           || | | |
Db        218 GVLSFTHHQIIELARDCLDKSHQGLITSRYFLELQHKLDKLL----------QEAHDRSE 267

Qy         88 SHSYLPELGKSEMEFIESKRPRLELLPDPLLRPSPLLATGQPAGSEDLTKDRSLTGKLEP 147
              |          |: ||:    :: ::   : ||: ||                ||
Db        268 S---------GELAFIKQLVRKILIV---IARPARLLEC-----------------LE- 296

Qy        148 VSPPSPPHTDPE-----LELVPPRLSKEELIQNMDRVDREITMVEQQISKLKKKQQQLEE 202
                     |||      ||      : : |:    : | |      ||:|   :  |||
Db        297 --------FDPEEFYYLLEAAEGHAKEGQGIKT--DIPRYI------ISQLGLNKDPLEE 340

Qy        203 EA---------AKPPEPEKPVSPPPIESKHRSLVQIIYDENRKKAEAAHRILEGLGPQVE 253
              |         |: || :: ||       | |           ||  |
Db        341 MAHLGNYDSGTAETPETDESVSSSNASLKLR----------RKPRE-------------- 376

Qy        254 LPLYNQPSDTRQYHENIKI--NQAMRKKLILYFKRRNHARKQWKQKFCQRY-----DQLM 306
                   ||     | ||:  | |      :|| |   :|:::  |   :      :|:
Db        377 -------SD----FETIKLISNGAYG---AVYFVRHKESRQRFAMKKINKQNLILRNQIQ 422

Qy        307 EA-LEKKVERIENNPRRRAKESKVREYYEKQFPEIRKQRELQERMQSRVG------QRGS 359
              :| :|: :    ||          :    :  :| | |:  |      :      |
Db        423 QAFVERDILTFAENP----------FVVSMYCSFETRRHLCMVMEYVEGGDCATLMKNM 471

Qy        360 G-LSMSAARSEHEVSEIIDGLSEQENLEKQMRQLAVIPPMLYDADQQRIKFINMN----G 414
              | | :  ||     :| :  |     | |   | |   |  |        || :      |
Db        472 GPLPVDMARM--YFAETVLALEYLHNYGIVHRDLK--PDNLLVTSMGHIKLTDFGLSKVG 527

Qy        415 LMADPMKVYK----------DRQVMNMWSEQEKETFREKFMQHPKNFGLIASFLERKTV 463
              ||:    :|:          |:||         |    :    | ::  :   |    |
Db        528 LMSMTTNLYEGHIEKDAREFLDKQVCGTPEYIAPEVILRQGYGKPVDWWAMGIILYEFLV 587

Qy        464 AECVLYYYLT---------------KKNE----NYKSLV-----RRSYRRRGKSQQQQQ 498
                || ::  |              :|:|    : : |:     :    | |     :
Db        588 G-CVPFFGDTPEELFGQVISDEINWPEKDEAPPPDAQDLITLLLRQNPLERLGTGGAYEV 646

Qy        499 QQQQQQQQQQQQPMPRSSQEEKDEKEKEKEAEKEEEKPEVENDKEDLLKEKTDDTSGED- 557
              :| :  :      : |   |   : | | :    : : |  :  |    |: |||: ||
Db        647 KQHRFFRSLDWNSLLRQKAEFIPQLESEDDTSYFDTRSEKYHHME---TEEEDDTNDEDF 703

Qy        558 NDEKEAVAS---KGRKTANSQGRRKGRITRSMANEANSEEAITPQQSAELASME-LNESS 613
              | |    :|   :  |  :|      |||:: | |   |:::   :|  | | | |: ||
Db        704 NVEIRQFSSCSHRFSKVFSS----IDRITQNSAEE--KEDSVDKTKSTTLPSTETLSWSS 757

Qy        614 RWTE-EEMETAKKGLLEHGRNWSAIARMVGSKTVSQCKNFYFNYKKRQNLDEILQQHKLK 672
                ::| ::: |:     |  |                          :|    |  | :  :
Db        758 EYSEMQQLSTSNSSDTESNR----------------------HKLSSGL---LPKLAIS 791

Qy        673 MEKERNARRKKKKAPAAASEEAAFPPVVEDEEMEASGVSGNEEEMVEEAEALHASGNEV- 731
              | |::   :   |    ||   | :  :|        :  | |:  |  : :|   |
Db        792 TEGEQD---EAASCPGDPHEEPGKPALPPEE------CAQEEPEVTTPASTISSSTLSVG 842

Qy        732 ----------PRGECSGPATVNNSSDTESIPSPH--------TEAAKDTGQNGPKPPATL 773
                       | ||    : :|||   | |: |         ||  | :|    |   :|
Db        843 SFSEHLDQINGRSECVD--STDNSSKPSSEPASHMARQRLESTEKKKISG----KVTKSL 896

Qy        774 GADG------------PPPGPPTPPRRTSRAPIEPTPASEATGAPTPPPAPPSPSAPPPV 821
              |                | | |  |   |        |:|    :|:   :  | |   |:
Db        897 SASALSLMIPGDMFAVSPLGSPMSPHSLS-----SDPSSSRDSSPSRDSSAASASPHQPI 951

Qy        822 VPKEEKEEE--TAAAPPVEEGE-------------EQKPPAAE----------------- 849
              |    :   |  |  |  |:            |:  ||  :
Db        952 VIHSSGKNYGFTIRAIRVYVGDSDIYTVHHIVWNVEEGSPACQAGLKAGDLITHINGEPV 1011

Qy        850 ---------ELAVDTGK----AEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEKK 896
                        || : :|       | ::      : |||:    ::       : |::||
Db       1012 HGLVHTEVIELLLKSGNKVSITTTPFEN---TSIKTGPARRNSYKS--RMVRRSKKSKKK 1066

Qy        897 EGGSGRAT----TAKSSGAPQDSDSSATCSADEVDEAEGGDKNRLLSPRPSLLTPTGDPR 952
              |    | :     ||       :  | :|              || ||   ||    | |
Db       1067 ESLERRRSLFKKLAKQPSPLLHTSRSFSC-----------LNRSLSSGESL---PGSPT 1111

Qy        953 ANASPQKPLDLKQLKQRAAAIPPIQVTKVHEPPREDAAPTKPAPPAPPPPQNLQPESDAP 1012
              : ||: |                       |   : |  |:        |:  | | ||
Db       1112 HSLSPRSP----------------------TPSYRSTPDFPS--GTNSSQSSSPSSSAP 1146

Qy       1013 QQPGSSPRGKSRSP-----APPADKEAFAAEAQKLPGDPPCWTSGLPFPVPPREVIKASP 1067
              |  |  |  |       ||    : : :  :|  |: |                 ||
Db       1147 NSPAGS--GHIRPSTLHGLAPKLGGQRYRSGRRKSAGNIP----------------LSP 1187

Qy       1068 HA--PDPSAFSYAPPGHPLPLGLHDTARPVLPR--PPTISNPPPLISSAKHPSVLERQIG 1123
              |  | |:    :|   | ||  |       : :  |  : :|| ::
Db       1188 LARTPSPTPQPTSPQRSPSPLLGHSLGNSKIAQAFPSKMHSPPTIVR------------- 1234

Qy       1124 AISQGMSVQLHVPYSEHAKAPVGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPES 1183
                     |:   : |: |  |:    |       :||:|  |  :: |  |       |
Db       1235 ----------HIVRPKSAEPPRSPL---LKRVQSEEKLSPSYGSDKKHLCSR-----KHS 1276

Qy       1184 LGVPTAQEASVLRGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHGTPADVLYKGTI 1243
              | | |:|  | |  :    |  |: : :    |  :   :| :          | :
Db       1277 LEV-TQEE--VQREQSQREAPLQSLDENVCDVPPLSRARPVEQGCL----------KRPV 1323

Qy       1244 TRIIG-EDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPH 1302
              :| :| ::|    |||    | | |  |: :                    | ||
Db       1324 SRKVGRQESVDDLDR---DKL-KAKVVVK------------------KADGF------- 1353

Qy       1303 ETAAPKRTYDMMEGRVGRAISSASIEGLMGRAIPPERHSPHHLKEQHHIRGSITQGIPRS 1362
                                                ||:    |   :: |  ||  :
Db       1354 --------------------------------PEKQESH---QKSHGPGSDLENFALF 1376

Qy       1363 YVEAQEDYLRREAKLLKREGTPPPPPPSRDLTEAYKTQA-------LGPLKLKPAHEGLV 1415
              :|       || |:          | : :  :  :: :|          || |     |:
Db       1377 KLE------EREKKVY---------PKAVERSSTFENKASMQEAPPLGSLLKDALHKQAS 1421

Qy       1416 ATVKEAGRSIHEIPREELRHTPELPLAPRPLKEGSITQGTPLKYDTGASTTG---SKKHD 1472
                |    :| |  :   :   || |   |:: |      | | | |    |
Db       1422 VRASEGAMSDGPVPAEHRQGGGDFRRAPAP---GTLQDGLCHSLDRGISGKGEGTEKSSQ 1478

Qy       1473 VRSLIGSPGRTFPPVHPLDVMADARALERACYEESLKSRPGTASSSGGSIARGAPVIVPE 1532
              : |:    :   :  :| :    :||    |::|                    ||:
Db       1479 AKELLRCE-KLDSKLANIDYLRKKMSLEDK--EDNL----------------CPVL--- 1515

Qy       1533 LGKPRQSPLTYEDHGAPFAGHLPRGSPVT----MREPTPRLQEGS-LSSSKASQDRKLTS 1587
                ||: :  ::|         || |:||      :|| |  :  : :||: |:|   ::
Db       1516 --KPKMTAGSHE--------CLP-GNPVRPTGGQQEPPPASESRAFVSSTHAAQMSAVSF 1564
```

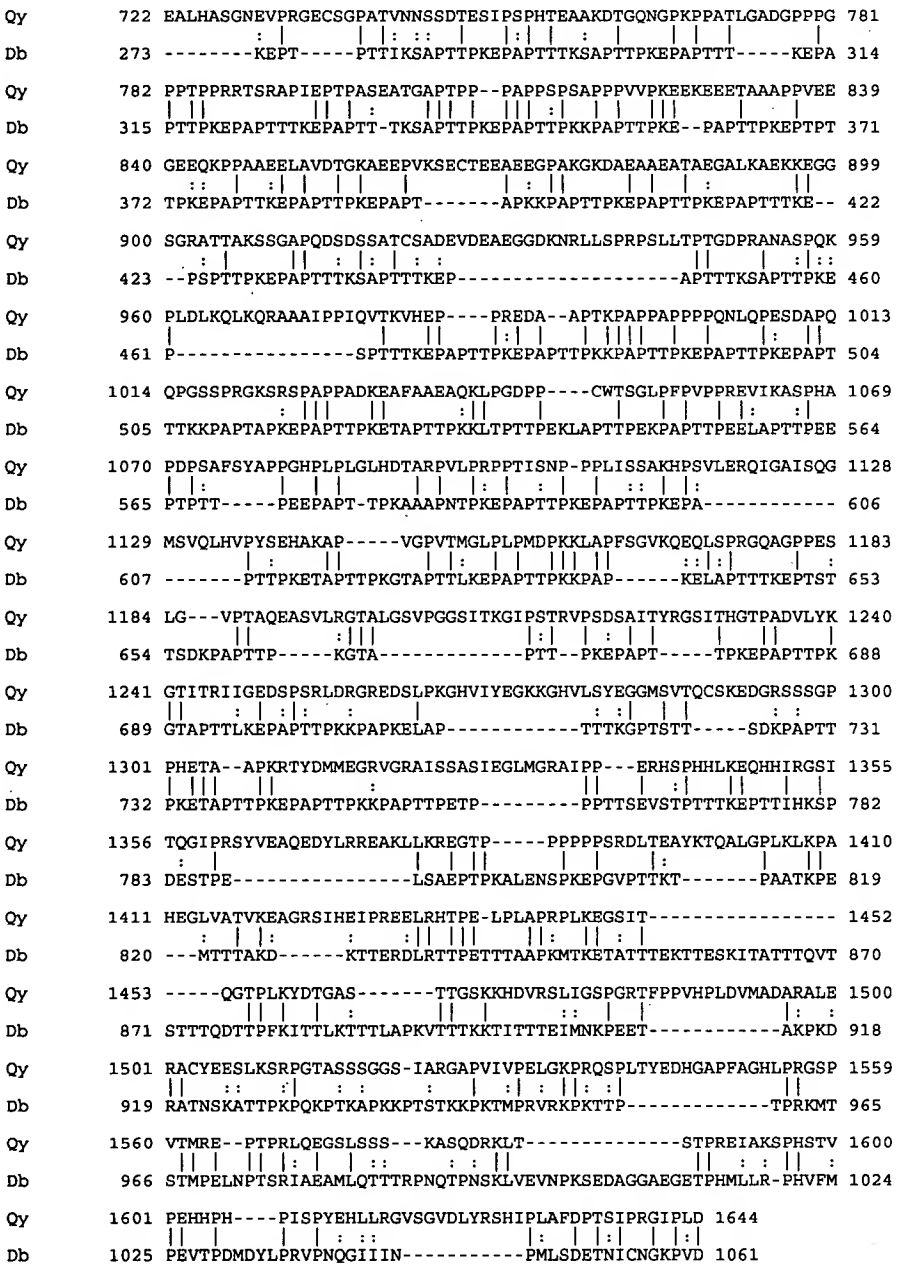
```
Qy       1588 TPREI--------AKSPHSTVPEHHPHPISPYEHLLRG-----------------VSGVD 1622
Db       1565 VPLKALTGRVDSGTEKPGLVAPE-SPVRKSPSEYKLEGRSVSCLEPIEGTLDIALLSGPQ 1623

Qy       1623 LYRSHIP---LAFDPT-------SIPRGIPL------DAAAAYYL-PRHLAPNPTYPHLY 1665
Db       1624 ASKTELPSPESAQSPSPSGDVRASVPPVLPSSSGKKNDTTSARELSPSSLKMNKSY--LL 1681

Qy       1666 PPYLIRGYPDTAALENRQTII---NDYITSQQMHHNTA----TAMAQRADMLRGLSPRES 1718
Db       1682 EPWFL---PPSRGLQNSPAVSLPDPEFKRDRKGPHPTARSPGTVMESNPQQREGSSPKHQ 1738

Qy       1719 SLALNYAAGPRGIIDLSQVPHLPVLVPPTPGTPATAMDRLAYLPTAPQPFSSRHSSSPLS 1778
Db       1739 ----DHTTDPKLLTCLGQNLHSPDLARP---------------------------RCPLP 1767

Qy       1779 PGGPTHLTKPTTTSSSER-------ERDRDR--------------ERDRDREREKSIL-- 1815
Db       1768 PEASPSREKPGLRESSERGPPTARSERSAARADTCREPSMELCFPETAKTSDNSKNLLSV 1827

Qy       1816 --------TSTTTVEHAPIWRPGTEQSSGSSGSSGGGGGSSSRPASHSHAHQHSPISPRT 1867
Db       1828 GRTHPDFYTQTQAMEKA--WAPG-----GKTNHKDGPG--EARPPPRDNSSLHSAGIPCE 1878

Qy       1868 QDALQQRPSVLHNTGMKGIITAVEPSKPTVLRSTSTSSPVRPAATFPPATHCPLGGTLDG 1927
Db       1879 KELGKVR-------------RGVEPKPEALLARRSLQ---------PPGIESEKSEKLSS 1916

Qy       1928 VYPTLM-----EPVLLPKEAPRVARPER-PRADTGHAFLAKPPAR---------SGLEPA 1972
Db       1917 -FPSLQKDGAKEPE--RKEQPLQRHPSSIPPPPLTAKDLSSPAARQHCSSPSHASGREPG 1973

Qy       1973 SSPS---KGSEPRPLVPPVSGHATIARTPAKNLAPHHASPDPPAPPASASDPHREKTQSK 2029
Db       1974 AKPSTAEPSSSPQDPPKPVAAHS--------ESSSHKPRPGPDPGPPKTKHPDRSLSSQK 2025

Qy       2030 PFSI-----QELELRSLG---YHGSSYSPEG--VEPVSPVS----SPSLTHDKGLPKHLE 2075
Db       2026 P-SVGATKGKEPATQSLGGSSREGKGHSKSGPDVFPATPGSQNKASDGIGQGEGGPSVPL 2084

Qy       2076 ELDKSHLEGELRPKQPG-PVKLGGEAAHLPHLRPLPESQPSSSPLLQTAPGVKGHQRVVT 2134
Db       2085 HTDRAPLDAKPQPTSGGRPLEVLEKPVHLPRPGHPGPSEPADQKL--SAVGEK------- 2135

Qy       2135 LAQHISEVITQDYTRHHPQQLSAPLPAPLYSFPGASCPVLDLR---RPPSDLYLPPPDHG 2191
Db       2136 ----------QTLSPKHPK------PSTVKDCP-TLCKQTDNRQTDKSPSQ--------- 2169

Qy       2192 APARGSPHSEGGKRSPE----PNKTSVLGGGEDGIEPVSPPEGMTEPGHSRSAVYPLLYR 2247
Db       2170 -PAANTDRRAEGKKCTEALYAPAEGDKLEAGLSFVHSENRLKGAERPAAGVGKGFP---- 2224

Qy       2248 DGEQTEPSRMGSKSPGNTSQPPAFFSKLTESNSAMVKSKKQEINKKLNTHNRNEPEYNIS 2307
Db       2225 --------EARGKGPG-PQKPP------TEAD-------------KPNGMKRSP------ 2250

Qy       2308 QPGTEIFNMPAITGTGLMTYRSQAVQEH----ASTNMGLEAIIRKALMGKYDQWEESPPL 2363
Db       2251 ------------SATGQSSFRSTALPEKSLSCSSSFPETRAGVREASAASSD-------- 2290

Qy       2364 SANAFNPLNASASLPAAMPITAADGRSDHTLTSPGGGGK---AKVSGRPSSR-------- 2412
Db       2291 --------TSSAKAAGGMLELPAPSNRDHRKAQPAGEGRTHMTKSDSLPSFRVSTLPLES 2342

Qy       2413 --------------------------KAKSPAPGLASGDRPPSVSSVHSEGDCNRRTPL 2445
Db       2343 HHPDPNTMGGASHRDRALSVTATVGETKGKDPAPA-----QPPPARKQNVGRDVTKPSPA 2397

Qy       2446 TNRVWEDRPSS 2456
Db       2398 PN---TDRPIS 2405
```

```
RESULT 38
US-10-124-557-104
; Sequence 104, Application US/10124557
; Publication No. US20020137894A1
;   GENERAL INFORMATION:
;        APPLICANT: Turner, Katherine
;                   Clark, Stephen C.
;                   Jacobs, Kenneth
;                   Hewick, Rodney M.
;                   Gesner, Thomas G.
;        TITLE OF INVENTION: Megakaryocyte Stimulating Factors
;        NUMBER OF SEQUENCES: 143
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Genetics Institute, Inc.
;             STREET: 87 CambridgePark Drive
;             CITY: Cambridge
;             STATE: Massachusetts
;             COUNTRY: U.S.A.
;             ZIP: 02140
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/124,557
;             FILING DATE: 16-Apr-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US 07/643,502
;             FILING DATE: 18-JAN-1991
;             APPLICATION NUMBER: US 07/546,114
;             FILING DATE: 29-JUN-1990
;             APPLICATION NUMBER: US 07/457,196
;             FILING DATE: 29-DEC-1989
;             APPLICATION NUMBER: US 07/390,901
;             FILING DATE: 08-AUG-1989
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Cserr, Luann
;             REGISTRATION NUMBER: 31,822
;             REFERENCE/DOCKET NUMBER: GI 5190
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (617)876-1170
;             TELEFAX: (617)876-5851
;   INFORMATION FOR SEQ ID NO: 104:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1140 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 104:
US-10-124-557-104

  Query Match             3.2%;  Score 423;  DB 13;  Length 1140;
  Best Local Similarity   21.1%;  Pred. No. 3.4e-11;
  Matches  269;  Conservative  133;  Mismatches  512;  Indels  360;  Gaps   57;

Qy        476 NENYKSLVRRSYRRRGKSQ----------QQQQQQQQQQQQQQQQPMPRSSQEEKDEKE 524
Db        126 SQTIKSTTKRSPKPPNKKKTKKVIESEEITEEHSVSENQESSSSSSSSSSSSTIWKIKSS 185

Qy        525 KEKEAEKE-EEKPEVENDKEDLLKEK-------TDDT-SGEDNDEKEAVASKGRKTANSQ 575
Db        186 KNSAANRELQKKLKVKDNKKNRTKKKPTPKPPVVDEAGSGLDNGDFKVTTPDTSTTQHNK 245

Qy        576 GRRKGRITRSMANEANSEEAITP-QQSAELASMELNESSRWTEEEMETAKKGLLEHGRNW 634
Db        246 VSTSPKIT--TAKPINPRPSLPPNSDTSKETSLTVNKETTVETKETTTTNKQTSTDGKEK 303

Qy        635 SAIARMVGSKTVSQCKNFYFNYKKRQNLDEILQQHKLKME---KERNARRKKKKAPAAAS 691
Db        304 TTSAKETQSIEKTSAKDL-------APTSKVLAKPTPKAETTTKGPALTTPKEPTPTTPK 356
```



RESULT 39
US-10-124-557-142
; Sequence 142, Application US/10124557
; Publication No. US20020137894A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cserr, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)876-1170
; TELEFAX: (617)876-5851
; INFORMATION FOR SEQ ID NO: 142:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1313 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 142:
US-10-124-557-142

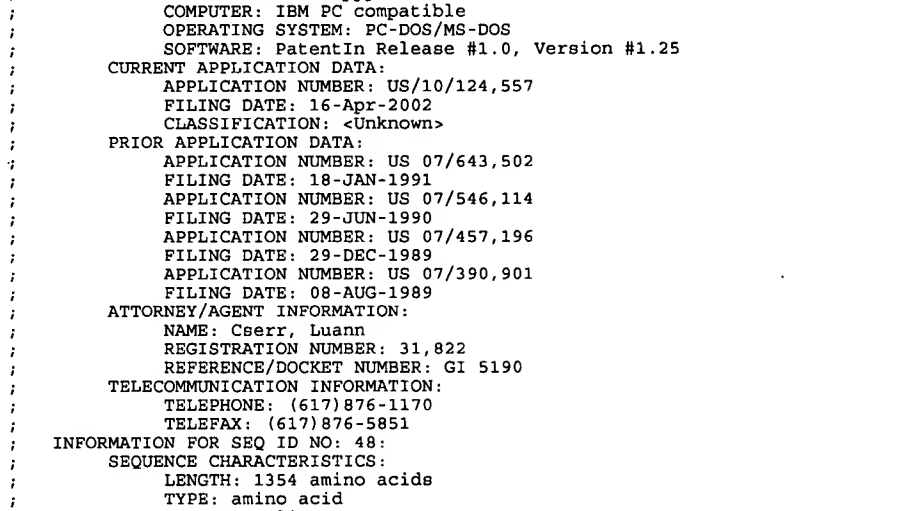
Query Match 3.2%; Score 421.5; DB 13; Length 1313;
Best Local Similarity 20.9%; Pred. No. 4.7e-11;
Matches 261; Conservative 131; Mismatches 507; Indels 349; Gaps 53;



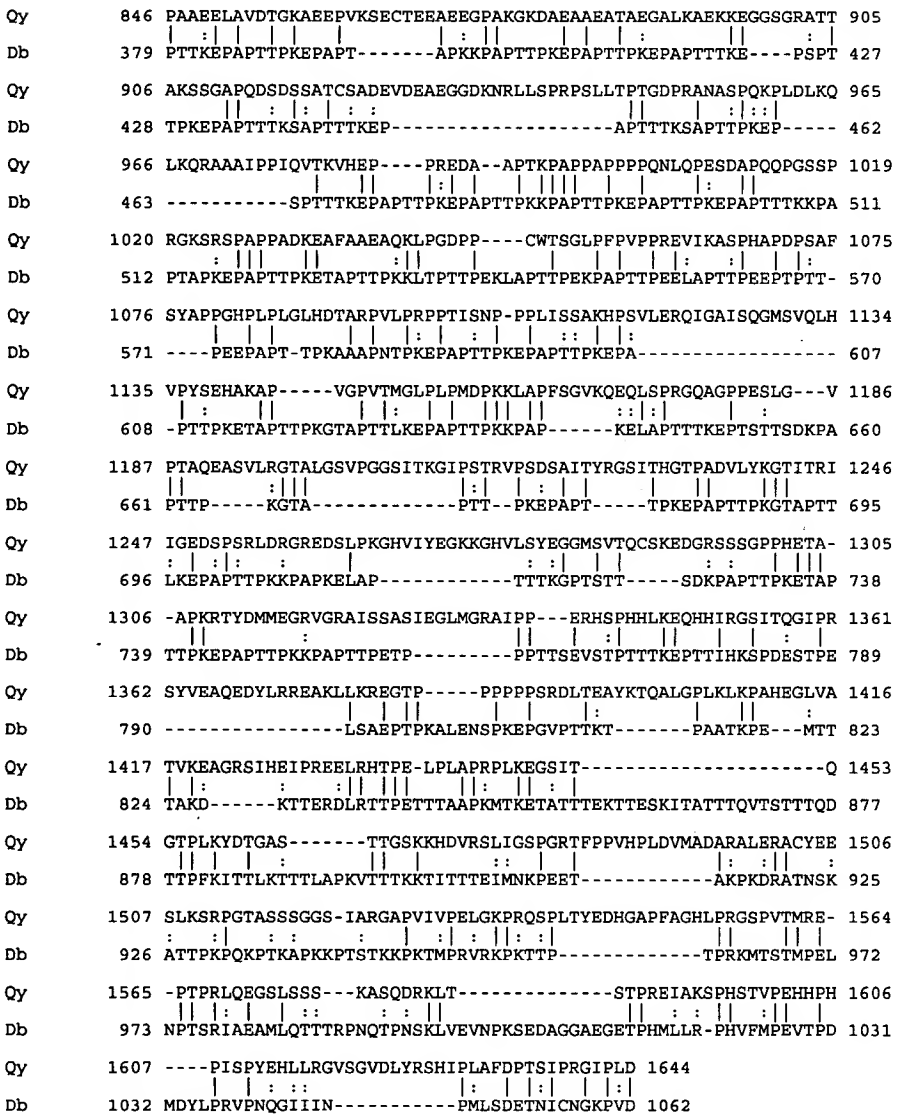
RESULT 40
US-10-124-557-48
; Sequence 48, Application US/10124557
; Publication No. US20020137894A1

; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cserr, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)876-1170
; TELEFAX: (617)876-5851
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1354 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 48:
US-10-124-557-48

Query Match 3.2%; Score 421.5; DB 13; Length 1354;
Best Local Similarity 20.9%; Pred. No. 4.8e-11;
Matches 261; Conservative 131; Mismatches 507; Indels 349; Gaps 53;

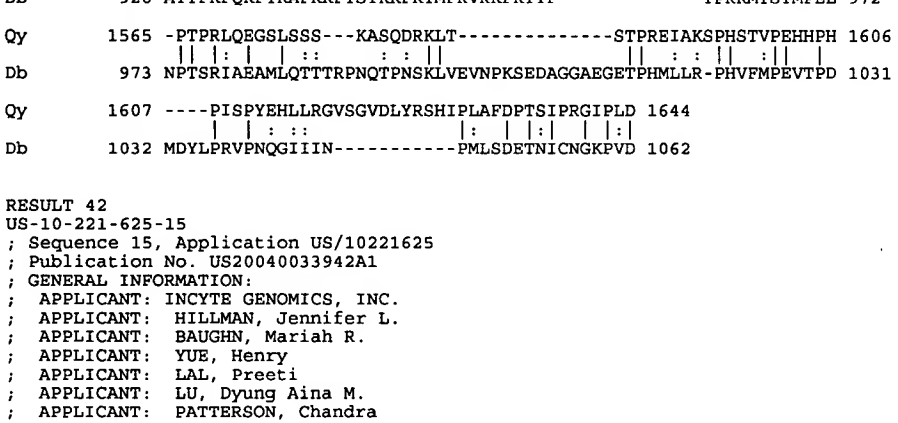


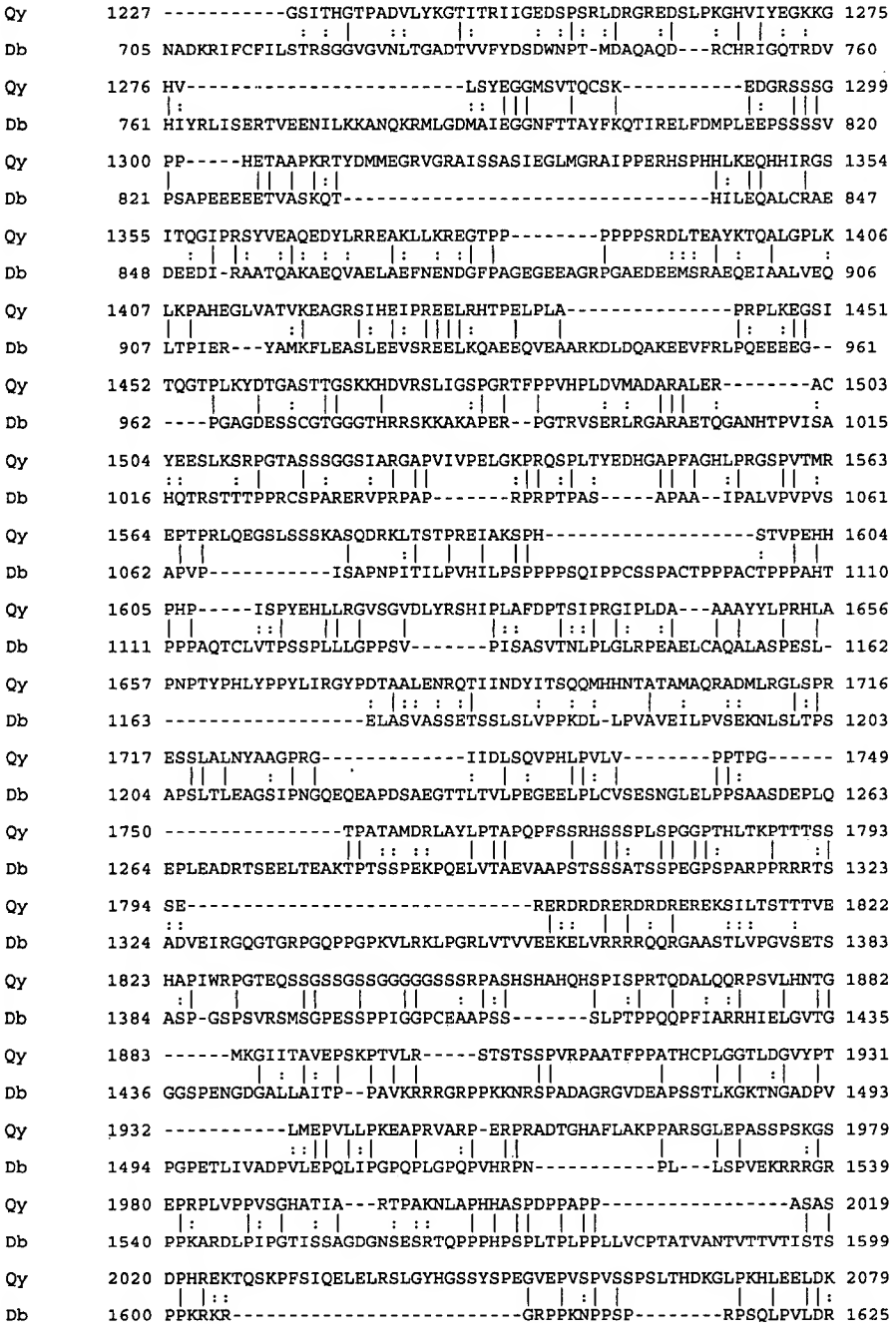
```
Qy        782 PPTPPRRTSRAPIEPTPASEATGAPTPP--PAPPSPSAPPPVVPKEEKEEETAAAPPVEE 839
Db        356 PTTPKEPAPTTTKEPAPTT-TKSAPTTPKEPAPTTPKKPAPTTPKE--PAPTTPKEPTPT 412

Qy        840 GEEQKPPAAEELAVDTGKAEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGG 899
Db        413 TPKEPAPTTKEPAPTTPKEPAPT-------APKKPAPTTPKEPAPTTPKEPAPTTTKE-- 463

Qy        900 SGRATTAKSSGAPQDSDSSATCSADEVDEAEGGDKNRLLSPRPSLLTPTGDPRANASPQK 959
Db        464 --PSPTTPKEPAPTTTKSAPTTTKEP--------------------APTTTKSAPTTPKE 501

Qy        960 PLDLKQLKQRAAAIPPIQVTKVHEP----PREDA--APTKPAPPAPPPPQNLQPESDAPQ 1013
Db        502 P---------------SPTTTKEPAPTTPKEPAPTTPKKPAPTTPKEPAPTTPKEPAPT 545

Qy       1014 QPGSSPRGKSRSPAPPADKEAFAAEAQKLPGDPP----CWTSGLPFPVPPREVIKASPHA 1069
Db        546 TTKKPAPTAPKEPAPTTPKETAPTTPKKLTPTTPEKLAPTTPEKPAPTTPEELAPTTPEE 605

Qy       1070 PDPSAFSYAPPGHPLPLGLHDTARPVLPRPPTISNP-PPLISSAKHPSVLERQIGAISQG 1128
Db        606 PTPTT-----PEEPAPT-TPKAAAPNTPKEPAPTTPKEPAPTTPKEPA------------ 647

Qy       1129 MSVQLHVPYSEHAKAP-----VGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPES 1183
Db        648 -------PTTPKETAPTTPKGTAPTTLKEPAPTTPKKPAP------KELAPTTTKEPTST 694

Qy       1184 LG---VPTAQEASVLRGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHGTPADVLYK 1240
Db        695 TSDKPAPTTP-----KGTA------------PTT--PKEPAPT-----TPKEPAPTTPK 729

Qy       1241 GTITRIIGEDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGP 1300
Db        730 GTAPTTLKEPAPTTPKKPAPKELAP-----------TTTKGPTSTT-----SDKPAPTT 772

Qy       1301 PHETA--APKRTYDMMEGRVGRAISSASIEGLMGRAIPP---ERHSPHHLKEQHHIRGSI 1355
Db        773 PKETAPTTPKEPAPTTPKKPAPTTPETP---------PPTTSEVSTPTTTKEPTTIHKSP 823

Qy       1356 TQGIPRSYVEAQEDYLRREAKLLKREGTP-----PPPPPSRDLTEAYKTQALGPLKLKPA 1410
Db        824 DESTPE---------------LSAEPTPKALENSPKEPGVPTTKT-------PAATKPE 860

Qy       1411 HEGLVATVKEAGRSIHEIPREELRHTPE-LPLAPRPLKEGSIT----------------- 1452
Db        861 ---MTTTAKD------KTTERDLRTTPETTTAAPKMTKETATTTEKTTESKITATTTQVT 911

Qy       1453 -----QGTPLKYDTGAS-------TTGSKKHDVRSLIGSPGRTFPPVHPLDVMADARALE 1500
Db        912 STTTQDTTPFKITTLKTTTLAPKVTTTKKTITTTEIMNKPEET-----------AKPKD 959

Qy       1501 RACYEESLKSRPGTASSSGGS-IARGAPVIVPELGKPRQSPLTYEDHGAPFAGHLPRGSP 1559
Db        960 RATNSKATTPKPQKPTKAPKKPTSTKKPKTMPRVRKPKTTP------------TPRKMT 1006

Qy       1560 VTMRE--PTPRLQEGSLSSS---KASQDRKLT--------------STPREIAKSPHSTV 1600
Db       1007 STMPELNPTSRIAEAMLQTTTRPNQTPNSKLVEVNPKSEDAGGAEGETPHMLLR-PHVFM 1065

Qy       1601 PEHHPH----PISPYEHLLRGVSGVDLYRSHIPLAFDPTSIPRGIPLD 1644
Db       1066 PEVTPDMDYLPRVPNQGIIIN----------PMLSDETNICNGKPVD 1102


RESULT 41
US-10-124-557-50
; Sequence 50, Application US/10124557
; Publication No. US20020137894A1
;   GENERAL INFORMATION:
;       APPLICANT: Turner, Katherine
;                  Clark, Stephen C.
;                  Jacobs, Kenneth
;                  Hewick, Rodney M.
;                  Gesner, Thomas G.
;       TITLE OF INVENTION: Megakaryocyte Stimulating Factors
;       NUMBER OF SEQUENCES: 143
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Genetics Institute, Inc.
;            STREET: 87 CambridgePark Drive
;            CITY: Cambridge
;            STATE: Massachusetts
;            COUNTRY: U.S.A.
;            ZIP: 02140
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/124,557
;            FILING DATE: 16-Apr-2002
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US 07/643,502
;            FILING DATE: 18-JAN-1991
;            APPLICATION NUMBER: US 07/546,114
;            FILING DATE: 29-JUN-1990
;            APPLICATION NUMBER: US 07/457,196
;            FILING DATE: 29-DEC-1989
;            APPLICATION NUMBER: US 07/390,901
;            FILING DATE: 08-AUG-1989
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Cserr, Luann
;            REGISTRATION NUMBER: 31,822
;            REFERENCE/DOCKET NUMBER: GI 5190
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (617)876-1170
;            TELEFAX: (617)876-5851
;   INFORMATION FOR SEQ ID NO: 50:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 1314 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 50:
US-10-124-557-50

  Query Match             3.2%;  Score 419.5;  DB 13;  Length 1314;
  Best Local Similarity   21.0%;  Pred. No. 5.8e-11;
  Matches  261;  Conservative  129;  Mismatches  503;  Indels  349;  Gaps   53;

Qy         501 QQQQQQQQQQPMPRSSQEEKDEKEKEKEAEKE-EEKPEVENDKEDLLKEK-------TDD 552
Db          72 ENQESSSSSSSSSSSSTIWKIKSSKNSAANRELQKKLKVKDNKKNRTKKKPTPKPPVVDE 131

Qy         553 T-SGEDNDEKEAVASKGRKTANSQGRRKGRITRSMANEANSEEAITP-QQSAELASMELN 610
Db         132 AGSGLDNGDFKVTTPDTSTTQHNKVSTSPKIT--TAKPINPRPSLPPNSDTSKETSLTVN 189

Qy         611 ESSRWTEEEMETAKKGLLEHGRNWSAIARMVGSKTVSQCKNFYFNYKKRQNLDEILQQHK 670
Db         190 KETTVETKETTTTNKQTSTDGKEKTTSAKETQSIEKTSAKDL-------APTSKVLAKPT 242

Qy         671 LKME---KERNARRKKKKAPAAASEEAAFPPVVEDEEMEASGVSGNEEEMVEEAEALHAS 727
Db         243 PKAETTTKGPALTTPKEPTPTTPKEPASTTP---------------------------- 273

Qy         728 GNEVPRGECSGPATVNNSSDTESIPSPHTEAAKDTGQNGPKPPATLGADGPPPGPPTPPR 787
Db         274 --KEPT-----PTTIKSAPTTPKEPAPTTTKSAPTTPKEPAPTTT-----KEPAPTTPKE 321

Qy         788 RTSRAPIEPTPASEATGAPTPP--PAPPSPSAPPPVVPKEEKEEETAAAPPVEEGEEQKP 845
Db         322 PAPTTTKEPAPTT-TKSAPTTPKEPAPTTPKKPAPTTPKE--PAPTTPKEPTPTTPKEPA 378
```

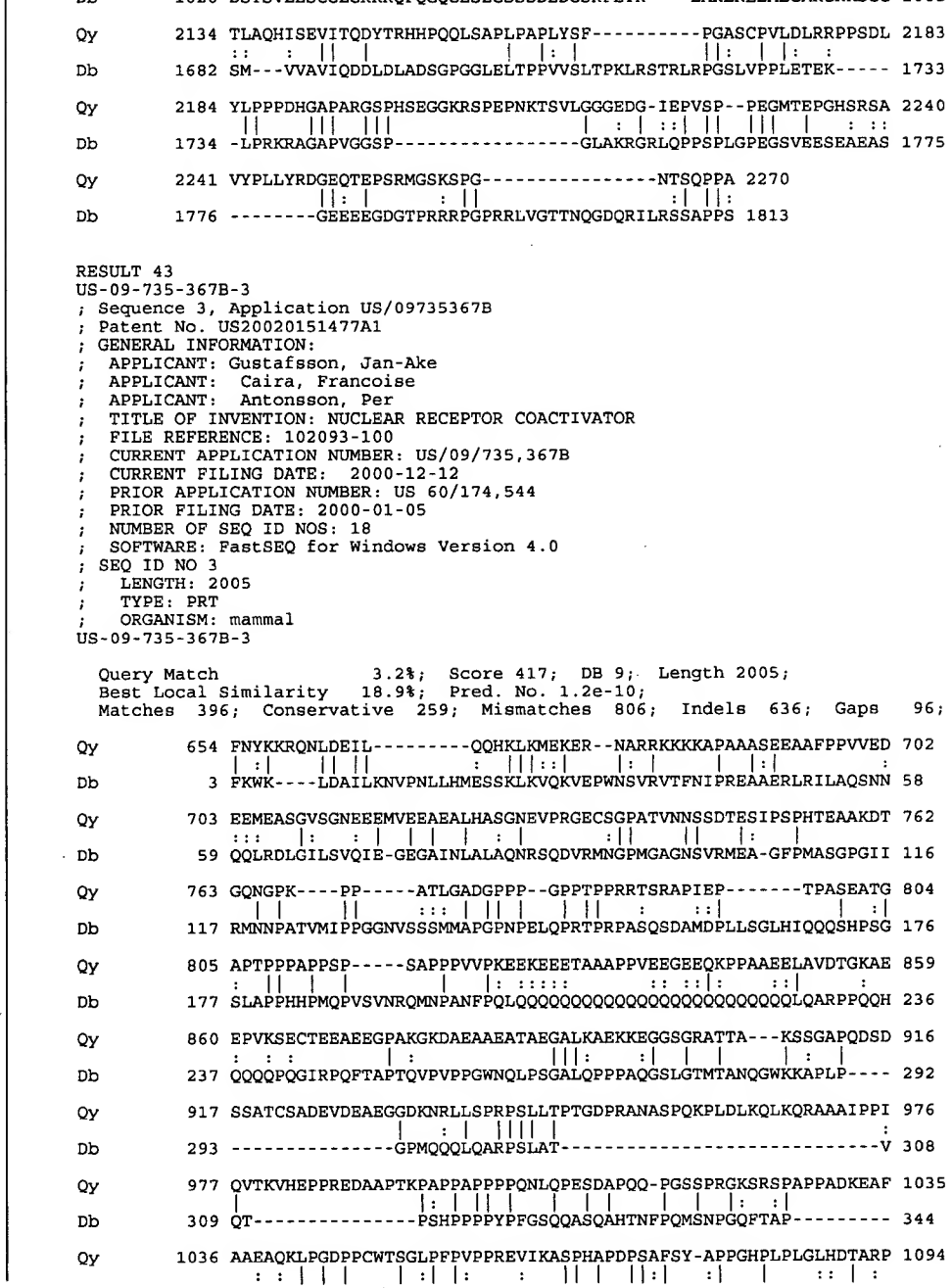


RESULT 42
US-10-221-625-15
; Sequence 15, Application US/10221625
; Publication No. US20040033942A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: YUE, Henry
; APPLICANT: LAL, Preeti
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: PATTERSON, Chandra
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: BANDMAN, Olga
; APPLICANT: TANG, Y. Tom
; APPLICANT: MATHUR, Preete
; APPLICANT: SHAH, Purvi
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: REDDY, Roopa
; TITLE OF INVENTION: TRANSCRIPTION FACTORS
; FILE REFERENCE: PF-0761 PCT
; CURRENT APPLICATION NUMBER: US/10/221,625
; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PERL Program
; SEQ ID NO 15
; LENGTH: 1828
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20040033942A1 879500CD1
US-10-221-625-15

Query Match 3.2%; Score 418; DB 12; Length 1828;
Best Local Similarity 19.8%; Pred. No. 1e-10;
Matches 402; Conservative 211; Mismatches 719; Indels 694; Gaps 80;







RESULT 43
US-09-735-367B-3
; Sequence 3, Application US/09735367B
; Patent No. US20020151477A1
; GENERAL INFORMATION:
; APPLICANT: Gustafsson, Jan-Ake
; APPLICANT: Caira, Francoise
; APPLICANT: Antonsson, Per
; TITLE OF INVENTION: NUCLEAR RECEPTOR COACTIVATOR
; FILE REFERENCE: 102093-100
; CURRENT APPLICATION NUMBER: US/09/735,367B
; CURRENT FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 60/174,544
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2005
; TYPE: PRT
; ORGANISM: mammal
US-09-735-367B-3

Query Match 3.2%; Score 417; DB 9; Length 2005;
Best Local Similarity 18.9%; Pred. No. 1.2e-10;
Matches 396; Conservative 259; Mismatches 806; Indels 636; Gaps 96;

```
Db        345 --QMKSLQGGP----SRVPTPLQQPHLTNKSP-ASSPSSFQQGSPASSPT---VNQTQQQ 394

Qy       1095 VLPRPPTISNP------PPLISSAKHPSVLERQIGAISQGMSVQLHVPYSEHAKAPVGPV 1148

Db        395 MGPRPPQ-NNPLPQGFQQPVSSPGRNPMV---QQGNVPPNFMVMQQQPPNQ------GPQ 444

Qy       1149 TMGLPLPMDPKKLAP-FSGVKQEQLSPRGQAGPPESLG-VPTAQEASVLRGTALGSVPGG 1206

Db        445 SLHPGLGGMPKRLPPGFSA---------GQANPNFMQGQVP----------STTATTPGN 485

Qy       1207 SITKGIPSTRVPSDSAITYRGSITHGTPADVLYKGTITRIIGEDSPSRLDRGREDSLPKG 1266

Db        486 S---GAP--QLQANQNVQHAGGQGAGPPQNQM--------------QVSHGPPNMMQPS 525

Qy       1267 HVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPHETAAPKRTYDMMEGRVGRAISSAS 1326

Db        526 LMGIHGNMNNQQAGTSGVPQVNLSNMQGQPQQGPP----------------------SQ 562

Qy       1327 IEGLMGRAIPPERHSPHHLKEQHHIRGSITQGIPRSYVEAQEDYLRREAKLL---KREGT 1383

Db        563 LMGMHQQIVP----SQGQMVQQ---QGTLNPQNPMILSRAQ---LMPQGQMMVNPPSQNL 612

Qy       1384 PPPP----PPSRDLTE------AYKTQALGP-----LKLKPAHEGLVATVKEAGRSIHEI 1428

Db        613 GPSPQRMTPPKQMLSQQGPQMMAPHNQMMGPQGQVLLQQNPMIEQIMTNQMQGNKQQFNT 672

Qy       1429 PREE--LRHTPELPLAPRPLKEGSITQGTPLKYDTGASTTGSKKHDVRSLIGSPGRTFPP 1486

Db        673 QNQSNVMPGPAQIMRGPTPNMQGNMVQFTGQMSGQMLPQQGPVNNSPSQVMGIQGQVLRP 732

Qy       1487 VHPLDVMADARALERACYEESLKSRPGTASSSGGSIARGAPVI-------VPELGKPRQ- 1538

Db        733 PGPSPHMAQQH------------GDPATTANNDVSLSQMMPDVSIQQTNMVPPHVQAMQG 780

Qy       1539 ---SPLTYEDHG----APFAG-------------HLPRGSPVTMREP-TPRLQEGSLSSS 1577

Db        781 NSASGNHFSGHGMSFNAPFSGAPNGNQMSCGQNPGFPVNKDVTLTSPLLVNLLQSDISAG 840

Qy       1578 KASQDRKLTST----------PREIAKSPHS-TVPEHHPHPISPYEH-LLRGVSGVDLYR 1625

Db        841 HFGVNNKQNNTNANKPKKKKPPRKKKNSQQDLNTPDTRPAGLEEADQPPLPGEQGISLDN 900

Qy       1626 SHIPLAFDPTSIPRGI---PLDAAAAYYLP----RHLAPNPTYPHLYP-PYLIR------ 1671

Db        901 SG-PKLPEFSNRPPGYPSQPVEQRPLQQMPPQLMQHVAPPPQPPQQQPQPQLPQQQQPPP 959

Qy       1672 -GYPDTAALENRQTIINDYITSQQMHHNTATAMAQRADMLRG-LSPRESSLALNYAAGPR 1729

Db        960 PSQPQSQQQQQQQQQMMMMLMMQQDPKSVRLPVSQNVHPPRGPLNPDSQRMPMQQSGSVP 1019

Qy       1730 GIIDLSQVPHLPVLVPPTP---------GTPATAMDRLAYLPTAPQ-PFSS--------- 1770

Db       1020 VMVSLQG----PASVPPSPDKQRMPMPVNTPLGSNSRKMVYQESPQNPSSSPLAEMASLP 1075

Qy       1771 --RHSSSPLSPGGPTHLTKPTTTSSSERERDRDRERDRDRDRDRDRDRDRDRDRDRDRDRD 1828

Db       1076 EASGSEAPSVPGGPNNMPSHV----------------------VLPQNQLMMTGP--K 1109

Qy       1829 PGTEQSSGSSGSSGGGGGSSSRPASHSHAHQHSP-ISPRTQDALQQRPSVLHNTGMKGII 1887

Db       1110 PGPSPLSATQGATPQQPPVNSLPSSHGH---HFPNVAAPTQTS---RPKTPNRASPRPYY 1163

Qy       1888 TAVEPSKPTVLRSTSTS-SPVRPAATFPPATHCPLGGTLDGVYPTLMEPVLLPKEAPRVA 1946

Db       1164 PQTPNNRPPSTEPSEISLSPER----------LNASIAGLFP----------PQIN 1199

Qy       1947 RPERPRADTGHAFLAKPPARSGLEPASSPSKGSEPRPLVPPVSGHATIARTPAKNLAPHH 2006

Db       1200 IPLPPRPNLNRGF-----DQQGLNPTTLKAIGQAPSNLTMNPSNFAT------------ 1241

Qy       2007 ASPDPPAPPASASDPHREKTQSKPFSIQELELRSLGYHGSSYSPEGVEPVSPVSSPSLTH 2066

Db       1242 -------------PQTHKLDSVVVN----------SGKQSNSGATKRASPSNS----- 1271

Qy       2067 DKGLPKHLEELDKSHLEGELRPKQPGPVKLGGEAAHLPHLRPLPESQPSSSPLLQTAPGV 2126

Db       1272 --------------------RRSSPGSSR---------KTTPSPGRQNSKAPKL------ 1296

Qy       2127 KGHQRVVTLAQHISEVITQDYTRHHPQQLSAPLPAPLYSFPGASCPVLDLRRPPSDLYLP 2186

Db       1297 -------TLASQTNAALLQ----------------------------NVELPRNVLVSP 1320

Qy       2187 PPDHGAPARGS-PHSEGGKRSPEPNKTSVLGGG---EDGIEPVSPP------------- 2228

Db       1321 TPLANPPVPGSFPNNSG--LNPQNSTVSVAAVGGVVEDNKESLNVPQDSDCQNSQSRKEQ 1378

Qy       2229 ---EGMTEPGHSRSAVYP--LLYRDGEQTEPSRMGSKSPGNTSQPPAFFSKLTESNSAMV 2283

Db       1379 VNIELKAVPAQEVKMVVPEDQSKKDGQPSDPNKLPS----------------VEENKNLV 1422

Qy       2284 KSKKQEINKKLN--THNRNEPEYNISQPG-TEI-FNMPAITGTGLM------TYRSQAVQ 2333

Db       1423 SPAMREAPTSLSQLLDNSGAPNVTIKPPGLTDLEVTPPVVSGEDLKKASVIPTLQDLSSS 1482

Qy       2334 EHASTNMGL-------EAIIRKALMGKYDQWEES-----------------PPLSANAF 2368

Db       1483 KEPSNSLNLPHSNELCSSLVHPELSEVSSNVAPSIPPVMSRPVSSSSISTPLPPNQITVF 1542

Qy       2369 ---NPLNASASLPAAMPITAADGRSDHTLTSPGGGGKAKVS---------GRP------- 2409

Db       1543 VTSNPITTSANTSAALPTHLQSALMSTVVTMPNAGSKVMVSEGQSAAQSNARPQFITPVF 1602

Qy       2410 -----------SSRKAKSPAPGLA--SGDRPPSVSSV---HSEGDCN-RRTPLTNRVWED 2452

Db       1603 INSSSIIQVMKGSQPSTIPAAPLTTNSGLMPPSVAVVGPLHIPQNIKFSSAPVP------ 1656

Qy       2453 RPSSAGSTPFPY----NPLIMRLQAGVMASPPPPGLPAGSGPLAGPHHAWDEEPKPL 2505

Db       1657 -PNALSSSPAPNIQTGRPLVLSSRATPVQLPSP---PCTSSPVV-PSHPPVQQVKEL 1708
```

```
RESULT 44
US-10-124-557-58
; Sequence 58, Application US/10124557
; Publication No. US20020137894A1
;   GENERAL INFORMATION:
;        APPLICANT: Turner, Katherine
;                   Clark, Stephen C.
;                   Jacobs, Kenneth
;                   Hewick, Rodney M.
;                   Gesner, Thomas G.
;        TITLE OF INVENTION: Megakaryocyte Stimulating Factors
;        NUMBER OF SEQUENCES: 143
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Genetics Institute, Inc.
;             STREET: 87 CambridgePark Drive
;             CITY: Cambridge
;             STATE: Massachusetts
;             COUNTRY: U.S.A.
;             ZIP: 02140
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/124,557
;             FILING DATE: 16-Apr-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US 07/643,502
;             FILING DATE: 18-JAN-1991
;             APPLICATION NUMBER: US 07/546,114
;             FILING DATE: 29-JUN-1990
;             APPLICATION NUMBER: US 07/457,196
;             FILING DATE: 29-DEC-1989
```

```
;             APPLICATION NUMBER: US 07/390,901
;             FILING DATE: 08-AUG-1989
;          ATTORNEY/AGENT INFORMATION:
;             NAME: Cserr, Luann
;             REGISTRATION NUMBER: 31,822
;             REFERENCE/DOCKET NUMBER: GI 5190
;          TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (617)876-1170
;             TELEFAX: (617)876-5851
;    INFORMATION FOR SEQ ID NO: 58:
;          SEQUENCE CHARACTERISTICS:
;             LENGTH: 1049 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;          MOLECULE TYPE: protein
;          SEQUENCE DESCRIPTION: SEQ ID NO: 58:
US-10-124-557-58

  Query Match             3.2%;  Score 416.5;  DB 13;  Length 1049;
  Best Local Similarity   21.2%;  Pred. No. 6.1e-11;
  Matches  264;  Conservative  128;  Mismatches  503;  Indels  349;  Gaps   56;

Qy        495 QQQQQQQQQQQQQQQQPMPRSSQEEKDEKEKEKEAEKE-EEKPEVENDKEDLLKEK---- 549
              ::    : |:           ||   | :  |   | :| ::| :|:::|::  |:|
Db         65 EEHSVSENQESSSSSSSSSSSSTIWKIKSSKNSAANRELQKKLKVKDNKKNRTKKKPTPK 124

Qy        550 ---TDDT-SGEDNDEKEAVASKGRKTANSQGRRKGRITRSMANEANSEEAITP-QQSAEL 604
                 |:  || || : :         | :::     :||   |   |   :: |   :::
Db        125 PPVVDEAGSGLDNGDFKVTTPDTSTTQHNKVSTSPKIT--TAKPINPRPSLPPNSDTSKE 182

Qy        605 ASMELNESSRWTEEEMETAKKGLLEHGRNWSAIARMVGSKTVSQCKNFYFNYKKRQNLDE 664
               |: :|: :    :|  |  |      |:  :  |:    |   :  |:            :
Db        183 TSLTVNKETTVETKETTTTNKQTSTDGKEKTTSAKETQSIEKTSAKDL-------APTSK 235

Qy        665 ILQQHKLKME---KERNARRKKKKAPAAASEEAAFPPVVEDEEMEASGVSGNEEEMVEEA 721
              :| :   | |   |       |:  |    | |:  |
Db        236 VLAKPTPKAETTTKGPALTTPKEPTPTTPKEPASTTP---------------------- 272

Qy        722 EALHASGNEVPRGECSGPATVNNSSDTESIPSPHTEAAKDTGQNGPKPPATLGADGPPPG 781
                     : |      | |: ::  |   |:| |  :  |    | |  |       |
Db        273 --------KEPT-----PTTIKSAPTTPKEPAPTTTKSAPTTPKEPAPTTT-----KEPA 314

Qy        782 PPTPPRRTSRAPIEPTPASEATGAPTPP--PAPPSPSAPPPVVPKEEKEEETAAAPPVEE 839
              | ||         || | :   ||| |  ||| :|  | |  |||      |     |
Db        315 PTTPKEPAPTTTKEPAPTT-TKSAPTTPKEPAPTTPKKPAPTTPKE--PAPTTPKEPTPT 371

Qy        840 GEEQKPPAAEELAVDTGKAEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGG 899
                ::  |  :| |  | |   |         | : ||     | |  | :       ||
Db        372 TPKEPAPTTKEPAPTTPKEPAPT-------APKKPAPTTPKEPAPTTPKEPAPTTTKE-- 422

Qy        900 SGRATTAKSSGAPQDSDSSATCSADEVDEAEGGDKNRLLSPRPSLLTPTGDPRANASPQK 959
                 : |    ||  : |: | : :                    ||    |  :|::
Db        423 --PSPTTPKEPAPTTTKSAPTTTKEP-------------------APTTTKSAPTTPKE 460

Qy        960 PLDLKQLKQRAAAIPPIQVTKVHEP----PREDA--APTKPAPPAPPPPQNLQPESDAPQ 1013
              |                   |   ||    |:| |   | ||||  |  |    |:  ||
Db        461 P---------------SPTTTKEPAPTTPKEPAPTTPKKPAPTTPKEPAPTTPKEPAPT 504

Qy       1014 QPGSSPRGKSRSPAPPADKEAFAAEAQKLPGDPP----CWTSGLPFPVPPREVIKASPHA 1069
                    : |||   ||      :||    |      |   | |  | |:   :|
Db        505 TTKKPAPTAPKEPAPTTPKETAPTTPKKLTPTTPEKLAPTTPEKPAPTTPEELAPTTPEE 564

Qy       1070 PDPSAFSYAPPGHPLPLGLHDTARPVLPRPPTISNP-PPLISSAKHPSVLERQIGAISQG 1128
              | |:      |  | |    | |  |: |  : |  |  :: | |:
Db        565 PTPTT-----PEEPAPT-TPKAAAPNTPKEPAPTTPKEPAPTTPKEPA------------ 606

Qy       1129 MSVQLHVPYSEHAKAP-----VGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPES 1183
                    | :    ||      | |:  | |  |||  ||      ::|:|     |  :
Db        607 -------PTTPKETAPTTPKGTAPTTLKEPAPTTPKKPAP------KELAPTTTKEPTST 653

Qy       1184 LG---VPTAQEASVLRGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHGTPADVLYK 1240
                 ||        :|||              |:|  | : | |      |   |'|    |
Db        654 TSDKPAPTTP-----KGTA-------------PTT--PKEPAPT-----TPKEPAPTTPK 688

Qy       1241 GTITRIIGEDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGP 1300
              ||    : | :|:   :      |                 : :|  | |       : :
Db        689 GTAPTTLKEPAPTTPKKPAPKELAP------------TTTKGPTSTT-----SDKPAPTT 731

Qy       1301 PHETA--APKRTYDMMEGRVGRAISSASIEGLMGRAIPP---ERHSPHHLKEQHHIRGSI 1355
              | |||   ||        :                  ||   | :|   ||   | |
Db        732 PKETAPTTPKEPAPTTPKKPAPTTPETP---------PPTTSEVSTPTTTKEPTTIHKSP 782

Qy       1356 TQGIPRSYVEAQEDYLRREAKLLKREGTP-----PPPPPSRDLTEAYKTQALGPLKLKPA 1410
               :  |                    |  | ||      |  |    |:         |  ||
Db        783 DESTPE---------------LSAEPTPKALENSPKEPGVPTTKT-------PAATKPE 819

Qy       1411 HEGLVATVKEAGRSIHEIPREELRHTPE-LPLAPRPLKEGSIT----------------- 1452
                 :  | |:     :    :|| |||    ||:  || : |
Db        820 ---MTTTAKD------KTTERDLRTTPETTTAAPKMTKETATTTEKTTESKITATTTQVT 870

Qy       1453 -----QGTPLKYDTGAS-------TTGSKKHDVRSLIGSPGRTFPPVHPLDVMADARALE 1500
                  || |  |  :        ||  |       ::  |  |            |: :
Db        871 STTTQDTTPFKITTLKTTTLAPKVTTTKKTITTTEIMNKPEET-----------AKPKD 918

Qy       1501 RACYEESLKSRPGTASSSGGS-IARGAPVIVPELGKPRQSPLTYEDHGAPFAGHLPRGSP 1559
              ||   ::   :|   : :     :   |  :| : ||: :|              ||
Db        919 RATNSKATTPKPQKPTKAPKKPTSTKKPKTMPRVRKPKTTP------------TPRKMT 965

Qy       1560 VTMRE--PTPRLQEGSLSSSKASQDRKLTSTPREIAKSPHSTVPEHHPHPISPYEHLLRG 1617
              || |  || |: |  | :          |: | :   :|:| : |      ::|      |
Db        966 STMPELNPTSRIAEAMLQT---------TTRPNQ---TPNSKLVE-----VNPKSEDAGG 1008

Qy       1618 VSGVD---LYRSHIPLAFDPTSIPRGIPLDAAAAYYLPRHLAPN 1658
                |     | | |:   | |  |    :|     ||||   ||
Db       1009 AEGETPHMLLRPHV---FMPEVTP---DMD-----YLPR--VPN 1039


RESULT 45
US-10-042-865-79
; Sequence 79, Application US/10042865
; Publication No. US20040029216A1
; GENERAL INFORMATION:
;  APPLICANT: Padigaru, Muralidhara
;  APPLICANT:  Li, Li
;  APPLICANT:  Zerhusen, Bryan D
;  APPLICANT:  Casman, Stacie J
;  APPLICANT:  Shenoy, Suresh G
;  APPLICANT:  Spytek, Kimberly
;  APPLICANT:  Zhong, Mei
;  APPLICANT:  Gangolli, Esha A
;  APPLICANT:  Burgess, Catherine E
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Vernet, Corine A.M
;  APPLICANT:  Taylor, Sarah
;  APPLICANT:  Tchernev, Velizar T
;  APPLICANT:  Miller, Charles E
;  APPLICANT:  Guo, Xiaojia
;  APPLICANT:  Boldog, Ference L
;  APPLICANT:  Grosse, William M
;  APPLICANT:  Alsobrook II, John P
;  APPLICANT:  Gerlach, Valerie L
;  APPLICANT:  Edinger, Shlomit R
;  APPLICANT:  Rothenberg, Mark E
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  MacDougall, John
;  APPLICANT:  Malyankar, Uriel M
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  Peyman, John
;  APPLICANT:  Smithson, Glennda
;  APPLICANT:  Gunther, Erik
;  APPLICANT:  Stone, David
;  TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
;  TITLE OF INVENTION:  Using the Same
;  FILE REFERENCE: 21402-537
;  CURRENT APPLICATION NUMBER: US/10/042,865
;  CURRENT FILING DATE:  2002-05-17
```

```
;  PRIOR APPLICATION NUMBER: 60/260,417
;  PRIOR FILING DATE: 2001-01-09
;  PRIOR APPLICATION NUMBER: 60/260,831
;  PRIOR FILING DATE: 2001-01-10
;  PRIOR APPLICATION NUMBER: 60/272,338
;  PRIOR FILING DATE: 2001-02-28
;  PRIOR APPLICATION NUMBER: 60/274,876
;  PRIOR FILING DATE: 2001-03-09
;  PRIOR APPLICATION NUMBER: 60/284,704
;  PRIOR FILING DATE: 2001-04-18
;  NUMBER OF SEQ ID NOS: 264
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 79
;   LENGTH: 2092
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-042-865-79

  Query Match             3.1%;  Score 414;  DB 12;  Length 2092;
  Best Local Similarity   20.0%;  Pred. No. 1.8e-10;
  Matches  452;  Conservative  245;  Mismatches  787;  Indels  780;  Gaps  112;

Qy        471 YLTKKNENYKSLVRRSYRRRGKSQQQQQQQQQQQQQQQQQPMPRSSQEE-----KDEKEK 525
Db        304 YEVKQHRFFRSLDWNSLLR-----------------QKAEFIPQLESEDDTSYFDTRSEK 346

Qy        526 EKEAEKEEEKPEVENDKEDLLKEKTDDTSGED-NDEKEAVAS---KGRKTANSQGRRKGR 581
Db        347 YHHMETEEE---------------DDTNDEDFNVEIRQFSSCSHRFSKVFSS----IDR 386

Qy        582 ITRSMANEANSEEAITPQQSAELASME-LNESSRWTE-EEMETAKKGLLEHGRNWSAIAR 639
Db        387 ITQNSAEE--KEDSVDKTKSTTLPSTETLSWSSEYSEMQQLSTSNSSDTESNR------- 437

Qy        640 MVGSKTVSQCKNFYFNYKKRQNLDEILQQHKLKMEKERNARRKKKKAPAAASEEAAFPPV 699
Db        438 ----------------HKLSSGL---LPKLAISTEGEQD---EAASCPGDPHEEPGKPAL 475

Qy        700 VEDEEMEASGVSGNEEEMVEEAEALHASGNEV-----------PRGECSGPATVNNSSDT 748
Db        476 PPEE------CAQEEPEVTTPASTISSSTLSVGSFSEHLDQINGRSECVD--STDNSSKP 527

Qy        749 ESIPSPH--------TEAAKDTGQNGPKPPATLGADG-----------PPPGPPTPPRR 788
Db        528 SSEPASHMARQRLESTEKKKISG----KVTKSLSASALSLMIPGDMFAVSPLGSPMSPHS 583

Qy        789 TSRAPIEPTPASEATGAPTPPPAPPSPSAPPPVVPKEEKEEE--TAAAPPVEEGE----- 841
Db        584 LS-----SDPSSSRDSSPSRDSSAASASPHQPIVIHSSGKNYGFTIRAIRVYVGDSDIYT 638

Qy        842 --------EQKPPAAE-------------------------ELAVDTGK----AEEPVK 863
Db        639 VHHIVWNVEEGSPACQAGLKAGDLITHINGEPVHGLVHTEVIELLLKSGNKVSITTTPFE 698

Qy        864 SECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGSGRAT----TAKSSGAPQDSDSSA 919
Db        699 N---TSIKTGPARRNSYKS--RMVRRSKKSKKKESLERRRSLFKKLAKQPSPLLHTSRSF 753

Qy        920 TCSADEVDEAEGGDKNRLLSPRPSLLTPTGDPRANASPQKPLDLKQLKQRAAAIPPIQVT 979
Db        754 SC------------LNRSLSSGESL---PGSPTHSLSPRSP------------------- 779

Qy        980 KVHEPPREDAAPTKPAPPAPPPPQNLQPESDAPQQPGSSPRGKSRSP-----APPADKEA 1034
Db        780 ----TPSYRSTPDFPS--GTNSSQSSSPSSSAPNSPAGS--GHIRPSTLHGLAPKLGGQR 831

Qy       1035 FAAEAQKLPGDPPCWTSGLPFPVPPREVIKASPHA--PDPSAFSYAPPGHPLPLGLHDTA 1092
Db        832 YRSGRRKSAGNIP----------------LSPLARTPSPTPQPTSPQRSPSPLLGHSLG 874

Qy       1093 RPVLPR--PPTISNPPPLISSAKHPSVLERQIGAISQGMSVQLHVPYSEHAKAPVGPVTM 1150
Db        875 NSKIAQAFPSKMHSPPTIVR---------------------HIVRPKSAEPPRSPL-- 909

Qy       1151 GLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPESLGVPTAQEASVLRGTALGSVPGGSITK 1210
Db        910 -LKRVQSEEKLSPSYGSDKKHLCSR-----KHSLEV-TQEE--VQREQSQREAPLQSLDE 960

Qy       1211 GIPSTRVPSDSAITYRGSITHGTPADVLYKGTITRIIG-EDSPSRLDRGREDSLPKGHVI 1269
Db        961 NVCDVPPLSRARPVEQGCL----------KRPVSRKVGRQESVDDLDR---DKL-KAKVV 1006

Qy       1270 YEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPHETAAPKRTYDMMEGRVGRAISSASIEG 1329
Db       1007 VK------------------KADGF---------------------------------- 1013

Qy       1330 LMGRAIPPERHSPHHLKEQHHIRGSITQGIPRSYVEAQEDYLRREAKLLKREGTPPPPPP 1389
Db       1014 -------PEKQESH---QKSHGPGSDLENFALFKLE------EREKKVY---------PK 1048

Qy       1390 SRDLTEAYKTQA-------LGPLKLKPAHEGLVATVKEAGRSIHEIPREELRHTPELPLA 1442
Db       1049 AVERSSTFENKASMQEAPPLGSLLKDALHKQASVRASEGAMSDGPVPAEHRQGGGDFRRA 1108

Qy       1443 PRPLKEGSITQGTPLKYDTGASTTG---SKKHDVRSLIGSPGRTFPPVHPLDVMADARAL 1499
Db       1109 PAP---GTLQDGLCHSLDRGISGKGEGTEKSSQAKELLRCE-KLDSKLANIDYLRKKMSL 1164

Qy       1500 ERACYEESLKSRPGTASSSGGSIARGAPVIVPELGKPRQSPLTYEDHGAPFAGHLPRGSP 1559
Db       1165 EDK--EDNL-----------------CPVL-----KPKMTAGSHE--------CLP-GNP 1191

Qy       1560 VT----MREPTPRLQEGS-LSSSKASQDRKLTSTPREI--------AKSPHSTVPEHHPH 1606
Db       1192 VRPTGGQQEPPPASESRAFVSSTHAAQMSAVSFVPLKALTGRVDSGTEKPGLVAPE-SPV 1250

Qy       1607 PISPYEHLLRG-----------------VSGVDLYRSHIP---LAFDPT-------SIPR 1639
Db       1251 RKSPSEYKLEGRSVSCLKPIEGTLDIALLSGPQASKTELPSPESAQSPSPSGDVRASVPP 1310

Qy       1640 GIPL------DAAAAYYL-PRHLAPNPTYPHLYPPYLIRGYPDTAALENRQTII---NDY 1689
Db       1311 VLPSSSGKKNDTTSARELSPSSLKMNKSY--LLEPWFL---PPSRGLQNSPAVSLPDPEF 1365

Qy       1690 ITSQQMHHNTA----TAMAQRADMLRGLSPRESSLALNYAAGPRGIIDLSQVPHLPVLVP 1745
Db       1366 KRDRKGPHPTARSPGTVMESNPQQREGSSPKHQ----DHTTDPKLLTCLGQNLHSPDLAR 1421

Qy       1746 PTPGTPATAMDRLAYLPTAPQPFSSRHSSSPLSPGGPTHLTKPTTTSSSER-------ER 1798
Db       1422 P--------------------------RCPLPPEASPSREKPGLRESSERGPPTARSER 1454

Qy       1799 DRDR--------------ERDRDREREKSIL----------TSTTTVEHAPIWRPGTEQS 1834
Db       1455 SAARADTCREPSMELCFPETAKTSDNSKNLLSVGRTHPDFYTQTQAMEKA--WAPG---- 1508

Qy       1835 SGSSGSSGGGGGSSSRPASHSHAHQHSPISPRTQDALQQRPSVLHNTGMKGIITAVEPSK 1894
Db       1509 -GKTNHKDGPG--EARPPPRDNSSLHSAGIPCEKELGKVR-------------RGVEPKP 1552

Qy       1895 PTVLRSTSTSSPVRPAATFPPATHCPLGGTLDGVYPTLM-----EPVLLPKEAPRVARPE 1949
Db       1553 EALLARRSLQ---------PPGIESEKSEKLSS-FPSLQKDGAKEPE--RKEQPLQRHPS 1600

Qy       1950 R-PRADTGHAFLAKPPAR---------SGLEPASSPS---KGSEPRPLVPPVSGHATIAR 1996
Db       1601 SIPPPPLTAKDLSSPAARQHCSSPSHASGREPGAKPSTAEPSSSPQDPPKPVAAHS---- 1656

Qy       1997 TPAKNLAPHHASPDPPAPPASASDPHREKTQSKPFSI-----QELELRSLG---YHGSSY 2048
Db       1657 ----ESSSHKPRPGPDPGPPKTKHPDRSLSSQKP-SVGATKGKEPATQSLGGSSREGKGH 1711

Qy       2049 SPEG--VEPVSPVS----SPSLTHDKGLPKHLEELDKSHLEGELRPKQPG-PVKLGGEAA 2101
Db       1712 SKSGPDVFPATPGSQNKASDGIGQGEGGPSVPLHTDRAPLDAKPQPTSGGRPLEVLEKPV 1771
```

```
Qy       2102 HLPHLRPLPESQPSSSPLLQTAPGVKGHQRVVTLAQHISEVITQDYTRHHPQQLSAPLPA 2161
Db       1772 HLPRPGHPGPSEPADQKL--SAVGEK----------------QTLSPKHPK------PS 1806

Qy       2162 PLYSFPGASCPVLDLR---RPPSDLYLPPPDHGAPARGSPHSEGGKRSPE----PNKTSV 2214
Db       1807 TVKDCP-TLCKQTDNRQTDKSPSQ----------PAANTDRRAEGKKCTEALYAPAEGDK 1855

Qy       2215 LGGGEDGIEPVSPPEGMTEPGHSRSAVYPLLYRDGEQTEPSRMGSKSPGNTSQPPAFFSK 2274
Db       1856 LEAGLSFVHSENRLKGAERPAAGVGKGFP----------EARGKGPG-PQKPP----- 1897

Qy       2275 LTESNSAMVKSKKQEINKKLNTHNRNEPEYNISQPGTEIFNMPAITGTGLMTYRSQAVQE 2334
Db       1898 -TEAD----------KPNGMKRSP----------------SATGQSSFRSTALPE 1925

Qy       2335 H----ASTNMGLEAIIRKALMGKYDQWEESPPLSANAFNPLNASASLPAAMPITAADGRS 2390
Db       1926 KSLSCSSSFPETRAGVREASAASSD----------------TSSAKAAGGMLELPAPSNR 1969

Qy       2391 DHTLTSPGGGGK---AKVSGRPSSR---------------------------------- 2412
Db       1970 DHRKAQPAGEGRTHMTKSDSLPSFRVSTLPLESHHPDPNTMGGASHRDRALSVTATVGET 2029

Qy       2413 KAKSPAPGLASGDRPPSVSSVHSEGDCNRRTPLTNRVWEDRPSS 2456
Db       2030 KGKDPAPA-----QPPPARKQNVGRDVTKPSPAPN---TDRPIS 2065
```

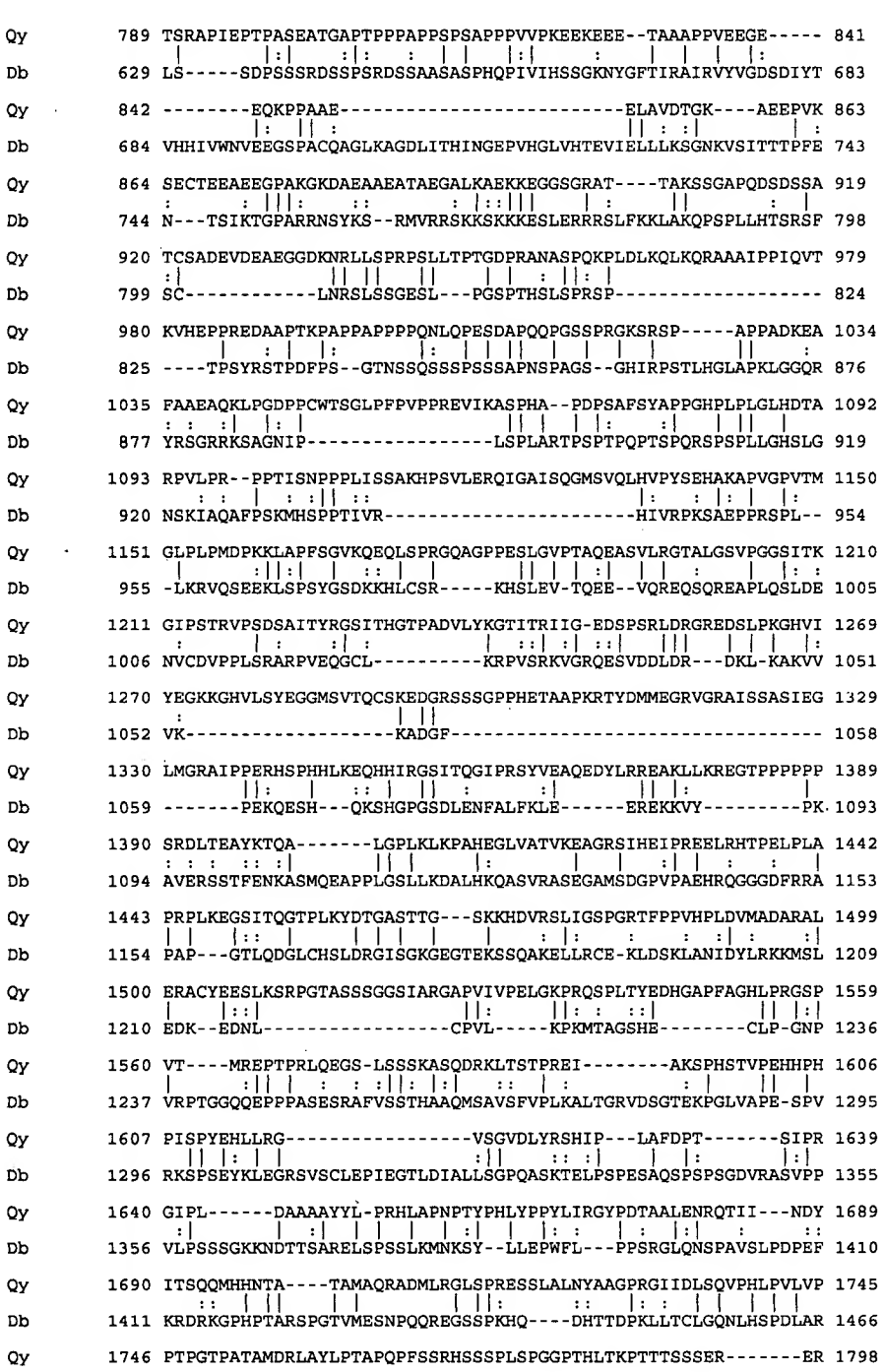
RESULT 46
US-10-377-035-18
; Sequence 18, Application US/10377035
; Publication No. US20030224406A1
; GENERAL INFORMATION:
; APPLICANT: Exelixis, Inc.
; TITLE OF INVENTION: MBCATs AS MODIFIERS OF THE BETA-CATENIN PATHWAY AND METHODS OF
; TITLE OF INVENTION: USE
; FILE REFERENCE: EX03-011C
; CURRENT APPLICATION NUMBER: US/10/377,035
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: 60/361,242
; PRIOR FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18
; LENGTH: 2092
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-377-035-18

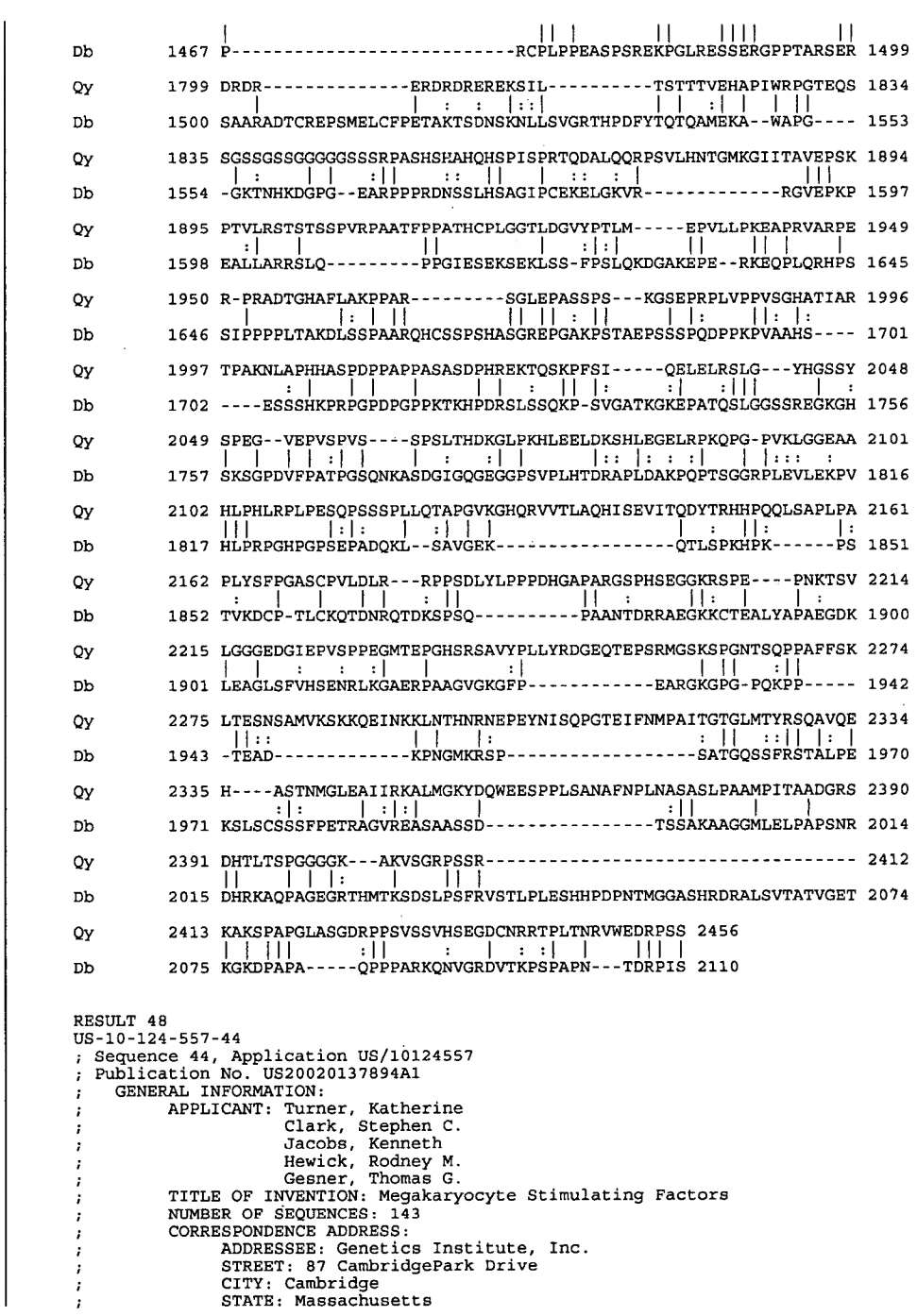
Query Match 3.1%; Score 414; DB 12; Length 2092;
Best Local Similarity 20.0%; Pred. No. 1.8e-10;
Matches 452; Conservative 245; Mismatches 787; Indels 780; Gaps 112;

```
Qy        471 YLTKKNENYKSLVRRSYRRRGKSQQQQQQQQQQQQQQQQQQPMPRSSQEE-----KDEKEK 525
Db        304 YEVKQHRFFRSLDWNSLLR----------------QKAEFIPQLESEDDTSYFDTRSEK 346

Qy        526 EKEAEKEEEKPEVENDKEDLLKEKTDDTSGED-NDEKEAVAS---KGRKTANSQGRRKGR 581
Db        347 YHHMETEEE---------------DDTNDEDFNVEIRQFSSCSHRFSKVFSS----IDR 386

Qy        582 ITRSMANEANSEEAITPQQSAELASME-LNESSRWTE-EEMETAKKGLLEHGRNWSAIAR 639
Db        387 ITQNSAEE--KEDSVDKTKSTTLPSTETLSWSSEYSEMQQLSTSNSSDTESNR------- 437

Qy        640 MVGSKTVSQCKNFYFNYKKRQNLDEILQQHKLKMEKERNARRKKKKAPAAASEEAAFPPV 699
Db        438 ---------------HKLSSGL---LPKLAISTEGEQD---EAASCPGDPHEEPGKPAL 475

Qy        700 VEDEEMEASGVSGNEEEMVEEAEALHASGNEV-----------PRGECSGPATVNNSSDT 748
Db        476 PPEE------CAQEEPEVTTPASTISSSTLSVGSFSEHLDQINGRSECVD--STDNSSKP 527

Qy        749 ESIPSPH--------TEAAKDTGQNGPKPPATLGADG-----------PPPGPPTPPRR 788
Db        528 SSEPASHMARQRLESTEKKKISG----KVTKSLSASALSLMIPGDMFAVSPLGSPMSPHS 583

Qy        789 TSRAPIEPTPASEATGAPTPPPAPPSPSAPPPVVPKEEKEEE--TAAAPPVEEGE----- 841
Db        584 LS-----SDPSSSRDSSPSRDSSAASASPHQPIVIHSSGKNYGFTIRAIRVYVGDSDIYT 638

Qy        842 --------EQKPPAAE-------------------------ELAVDTGK----AEEPVK 863
Db        639 VHHIVWNVEEGSPACQAGLKAGDLITHINGEPVHGLVHTEVIELLLKSGNKVSITTTPFE 698

Qy        864 SECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGSGRAT----TAKSSGAPQDSDSSA 919
Db        699 N---TSIKTGPARRNSYKS--RMVRRSKKSKKKESLERRRSLFKKLAKQPSPLLHTSRSF 753

Qy        920 TCSADEVDEAEGGDKNRLLSPRPSLLTPTGDPRANASPQKPLDLKQLKQRAAAIPPIQVT 979
Db        754 SC------------LNRSLSSGESL---PGSPTHSLSPRSP------------------- 779

Qy        980 KVHEPPREDAAPTKPAPPAPPPPQNLQPESDAPQQPGSSPRGKSRSP-----APPADKEA 1034
Db        780 ----TPSYRSTPDFPS--GTNSSQSSSPSSSAPNSPAGS--GHIRPSTLHGLAPKLGGQR 831

Qy       1035 FAAEAQKLPGDPPCWTSGLPFPVPPREVIKASPHA--PDPSAFSYAPPGHPLPLGLHDTA 1092
Db        832 YRSGRRKSAGNIP----------------LSPLARTPSPTPQPTSPQRSPSPLLGHSLG 874

Qy       1093 RPVLPR--PPTISNPPPLISSAKHPSVLERQIGAISQGMSVQLHVPYSEHAKAPVGPVTM 1150
Db        875 NSKIAQAFPSKMHSPPTIVR----------------------HIVRPKSAEPPRSPL-- 909

Qy       1151 GLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPESLGVPTAQEASVLRGTALGSVPGGSITK 1210
Db        910 -LKRVQSEEKLSPSYGSDKKHLCSR-----KHSLEV-TQEE--VQREQSQREAPLQSLDE 960

Qy       1211 GIPSTRVPSDSAITYRGSITHGTPADVLYKGTITRIIG-EDSPSRLDRGREDSLPKGHVI 1269
Db        961 NVCDVPPLSRARPVEQGCL----------KRPVSRKVGRQESVDDLDR---DKL-KAKVV 1006

Qy       1270 YEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPHETAAPKRTYDMMEGRVGRAISSASIEG 1329
Db       1007 VK------------------KADGF---------------------------------- 1013

Qy       1330 LMGRAIPPERHSPHHLKEQHHIRGSITQGIPRSYVEAQEDYLRREAKLLKREGTPPPPPP 1389
Db       1014 -------PEKQESH---QKSHGPGSDLENFALFKLE------EREKKVY---------PK 1048

Qy       1390 SRDLTEAYKTQA-------LGPLKLKPAHEGLVATVKEAGRSIHEIPREELRHTPELPLA 1442
Db       1049 AVERSSTFENKASMQEAPPLGSLLKDALHKQASVRASEGAMSDGPVPAEHRQGGGDFRRA 1108

Qy       1443 PRPLKEGSITQGTPLKYDTGASTTG---SKKHDVRSLIGSPGRTFPPVHPLDVMADARAL 1499
Db       1109 PAP---GTLQDGLCHSLDRGISGKGEGTEKSSQAKELLRCE-KLDSKLANIDYLRKKMSL 1164

Qy       1500 ERACYEESLKSRPGTASSSGGSIARGAPVIVPELGKPRQSPLTYEDHGAPFAGHLPRGSP 1559
Db       1165 EDK--EDNL----------------CPVL-----KPKMTAGSHE--------CLP-GNP 1191

Qy       1560 VT----MREPTPRLQEGS-LSSSKASQDRKLTSTPREI--------AKSPHSTVPEHHPH 1606
Db       1192 VRPTGGQQEPPPASESRAFVSSTHAAQMSAVSFVPLKALTGRVDSGTEKPGLVAPE-SPV 1250

Qy       1607 PISPYEHLLRG----------------VSGVDLYRSHIP---LAFDPT-------SIPR 1639
Db       1251 RKSPSEYKLEGRSVSCLKPIEGTLDIALLSGPQASKTELPSPESAQSPSPSGDVRASVPP 1310

Qy       1640 GIPL------DAAAAYYL-PRHLAPNPTYPHLYPPYLIRGYPDTAALENRQTII---NDY 1689
Db       1311 VLPSSSGKKNDTTSARELSPSSLKMNKSY--LLEPWFL---PPSRGLQNSPAVSLPDPEF 1365
```

```
Qy       1690 ITSQQMHHNTA----TAMAQRADMLRGLSPRESSLALNYAAGPRGIIDLSQVPHLPVLVP 1745
Db       1366 KRDRKGPHPTARSPGTVMESNPQQREGSSPKHQ----DHTTDPKLLTCLGQNLHSPDLAR 1421

Qy       1746 PTPGTPATAMDRLAYLPTAPQPFSSRHSSSPLSPGGPTHLTKPTTTSSSER-------ER 1798
Db       1422 P--------------------------RCPLPPEASPSREKPGLRESSERGPPTARSER 1454

Qy       1799 DRDR--------------ERDRDREREKSIL----------TSTTTVEHAPIWRPGTEQS 1834
Db       1455 SAARADTCREPSMELCFPETAKTSDNSKNLLSVGRTHPDFYTQTQAMEKA--WAPG---- 1508

Qy       1835 SGSSGSSGGGGGSSSRPASHSHAHQHSPISPRTQDALQQRPSVLHNTGMKGIITAVEPSK 1894
Db       1509 -GKTNHKDGPG--EARPPPRDNSSLHSAGIPCEKELGKVR------------RGVEPKP 1552

Qy       1895 PTVLRSTSTSSPVRPAATFPPATHCPLGGTLDGVYPTLM-----EPVLLPKEAPRVARPE 1949
Db       1553 EALLARRSLQ---------PPGIESEKSEKLSS-FPSLQKDGAKEPE--RKEQPLQRHPS 1600

Qy       1950 R-PRADTGHAFLAKPPAR---------SGLEPASSPS---KGSEPRPLVPPVSGHATIAR 1996
Db       1601 SIPPPPLTAKDLSSPAARQHCSSPSHASGREPGAKPSTAEPSSSPQDPPKPVAAHS---- 1656

Qy       1997 TPAKNLAPHHASPDPPAPPASASDPHREKTQSKPFSI-----QELELRSLG---YHGSSY 2048
Db       1657 ----ESSSHKPRPGPDPGPPKTKHPDRSLSSQKP-SVGATKGKEPATQSLGGSSREGKGH 1711

Qy       2049 SPEG--VEPVSPVS----SPSLTHDKGLPKHLEELDKSHLEGELRPKQPG-PVKLGGEAA 2101
Db       1712 SKSGPDVFPATPGSQNKASDGIGQGEGGPSVPLHTDRAPLDAKPQPTSGGRPLEVLEKPV 1771

Qy       2102 HLPHLRPLPESQPSSSPLLQTAPGVKGHQRVVTLAQHISEVITQDYTRHHPQQLSAPLPA 2161
Db       1772 HLPRPGHPGPSEPADQKL--SAVGEK-----------------QTLSPKHPK------PS 1806

Qy       2162 PLYSFPGASCPVLDLR---RPPSDLYLPPPDHGAPARGSPHSEGGKRSPE----PNKTSV 2214
Db       1807 TVKDCP-TLCKQTDNRQTDKSPSQ----------PAANTDRRAEGKKCTEALYAPAEGDK 1855

Qy       2215 LGGGEDGIEPVSPPEGMTEPGHSRSAVYPLLYRDGEQTEPSRMGSKSPGNTSQPPAFFSK 2274
Db       1856 LEAGLSFVHSENRLKGAERPAAGVGKGFP------------EARGKGPG-PQKPP----- 1897

Qy       2275 LTESNSAMVKSKKQEINKKLNTHNRNEPEYNISQPGTEIFNMPAITGTGLMTYRSQAVQE 2334
Db       1898 -TEAD-------------KPNGMKRSP-----------------SATGQSSFRSTALPE 1925

Qy       2335 H----ASTNMGLEAIIRKALMGKYDQWEESPPLSANAFNPLNASASLPAAMPITAADGRS 2390
Db       1926 KSLSCSSSFPETRAGVREASAASSD---------------TSSAKAAGGMLELPAPSNR 1969

Qy       2391 DHTLTSPGGGGK---AKVSGRPSSR---------------------------------- 2412
Db       1970 DHRKAQPAGEGRTHMTKSDSLPSFRVSTLPLESHHPDPNTMGGASHRDRALSVTATVGET 2029

Qy       2413 KAKSPAPGLASGDRPPSVSSVHSEGDCNRRTPLTNRVWEDRPSS 2456
Db       2030 KGKDPAPA-----QPPPARKQNVGRDVTKPSPAPN---TDRPIS 2065
```

```
RESULT 47
US-10-042-865-81
; Sequence 81, Application US/10042865
; Publication No. US20040029216A1
; GENERAL INFORMATION:
;  APPLICANT: Padigaru, Muralidhara
;  APPLICANT:  Li, Li
;  APPLICANT:  Zerhusen, Bryan D
;  APPLICANT:  Casman, Stacie J
;  APPLICANT:  Shenoy, Suresh G
;  APPLICANT:  Spytek, Kimberly
;  APPLICANT:  Zhong, Mei
;  APPLICANT:  Gangolli, Esha A
;  APPLICANT:  Burgess, Catherine E
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Vernet, Corine A.M
;  APPLICANT:  Taylor, Sarah
;  APPLICANT:  Tchernev, Velizar T
;  APPLICANT:  Miller, Charles E
;  APPLICANT:  Guo, Xiaojia
;  APPLICANT:  Boldog, Ference L
;  APPLICANT:  Grosse, William M
;  APPLICANT:  Alsobrook II, John P
;  APPLICANT:  Gerlach, Valerie L
;  APPLICANT:  Edinger, Shlomit R
;  APPLICANT:  Rothenberg, Mark E
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  MacDougall, John
;  APPLICANT:  Malyankar, Uriel M
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  Peyman, John
;  APPLICANT:  Smithson, Glennda
;  APPLICANT:  Gunther, Erik
;  APPLICANT:  Stone, David
;  TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
;  TITLE OF INVENTION:  Using the Same
;  FILE REFERENCE: 21402-537
;  CURRENT APPLICATION NUMBER: US/10/042,865
;  CURRENT FILING DATE:  2002-05-17
;  PRIOR APPLICATION NUMBER: 60/260,417
;  PRIOR FILING DATE: 2001-01-09
;  PRIOR APPLICATION NUMBER: 60/260,831
;  PRIOR FILING DATE: 2001-01-10
;  PRIOR APPLICATION NUMBER: 60/272,338
;  PRIOR FILING DATE: 2001-02-28
;  PRIOR APPLICATION NUMBER: 60/274,876
;  PRIOR FILING DATE: 2001-03-09
;  PRIOR APPLICATION NUMBER: 60/284,704
;  PRIOR FILING DATE: 2001-04-18
;  NUMBER OF SEQ ID NOS: 264
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 81
;   LENGTH: 2137
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-042-865-81

  Query Match             3.1%;  Score 414;  DB 12;  Length 2137;
  Best Local Similarity   20.0%;  Pred. No. 1.8e-10;
  Matches  452;  Conservative  245;  Mismatches  787;  Indels  780;  Gaps  112;

Qy        471 YLTKKNENYKSLVRRSYRRRGKSQQQQQQQQQQQQQQQQQPMPRSSQEE-----KDEKEK 525
Db        349 YEVKQHRFFRSLDWNSLLR----------------QKAEFIPQLESEDDTSYFDTRSEK 391

Qy        526 EKEAEKEEEKPEVENDKEDLLKEKTDDTSGED-NDEKEAVAS---KGRKTANSQGRRKGR 581
Db        392 YHHMETEEE---------------DDTNDEDFNVEIRQFSSCSHRFSKVFSS----IDR 431

Qy        582 ITRSMANEANSEEAITPQQSAELASME-LNESSRWTE-EEMETAKKGLLEHGRNWSAIAR 639
Db        432 ITQNSAEE--KEDSVDKTKSTTLPSTETLSWSSEYSEMQQLSTSNSSDTESNR------- 482

Qy        640 MVGSKTVSQCKNFYFNYKKRQNLDEILQQHKLKMEKERNARRKKKKAPAAASEEAAFPPV 699
Db        483 ---------------HKLSSGL---LPKLAISTEGEQD---EAASCPGDPHEEPGKPAL 520

Qy        700 VEDEEMEASGVSGNEEEMVEEAEALHASGNEV-----------PRGECSGPATVNNSSDT 748
Db        521 PPEE------CAQEEPEVTTPASTISSSTLSVGSFSEHLDQINGRSECVD--STDNSSKP 572

Qy        749 ESIPSPH--------TEAAKDTGQNGPKPPATLGADG-----------PPPGPPTPPRR 788
Db        573 SSEPASHMARQRLESTEKKKISG----KVTKSLSASALSLMIPGDMFAVSPLGSPMSPHS 628
```





RESULT 48
US-10-124-557-44
; Sequence 44, Application US/10124557
; Publication No. US20020137894A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts

```
;            COUNTRY: U.S.A.
;            ZIP: 02140
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/124,557
;            FILING DATE: 16-Apr-2002
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US 07/643,502
;            FILING DATE: 18-JAN-1991
;            APPLICATION NUMBER: US 07/546,114
;            FILING DATE: 29-JUN-1990
;            APPLICATION NUMBER: US 07/457,196
;            FILING DATE: 29-DEC-1989
;            APPLICATION NUMBER: US 07/390,901
;            FILING DATE: 08-AUG-1989
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Cserr, Luann
;            REGISTRATION NUMBER: 31,822
;            REFERENCE/DOCKET NUMBER: GI 5190
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (617)876-1170
;            TELEFAX: (617)876-5851
;   INFORMATION FOR SEQ ID NO: 44:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 1270 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 44:
US-10-124-557-44

  Query Match             3.1%;  Score 410;  DB 13;  Length 1270;
  Best Local Similarity   21.0%;  Pred. No. 1.5e-10;
  Matches  262;  Conservative  128;  Mismatches  489;  Indels  368;  Gaps   54;

Qy        515 SSQE--------EKDEKEKEKEAEKEEEK------------PEVENDKEDLLKEK----- 549
              ||||        |  |: :| : : : :|              ||:::|::  |:|
Db         23 SSQELSCKGRCFESFERGRECDCDAQCKKYDKCCPDYESFCAEVKDNKKNRTKKKPTPKP 82

Qy        550 --TDDT-SGEDNDEKEAVASKGRKTANSQGRRKGRITRSMANEANSEEAITP-QQSAELA 605
                |:  || || : :       | :::    :||   |   |   :: |   :::
Db         83 PVVDEAGSGLDNGDFKVTTPDTSTTQHNKVSTSPKIT--TAKPINPRPSLPPNSDTSKET 140

Qy        606 SMELNESSRWTEEEMETAKKGLLEHGRNWSAIARMVGSKTVSQCKNFYFNYKKRQNLDEI 665
              |: :|: :    :|  |  |     |:  :  |:   |   :  |:           ::
Db        141 SLTVNKETTVETKETTTTNKQTSTDGKEKTTSAKETQSIEKTSAKDL-------APTSKV 193

Qy        666 LQQHKLKME---KERNARRKKKKAPAAASEEAAFPPVVEDEEMEASGVSGNEEEMVEEAE 722
              | :   | |   |       |:  |    | |:  |
Db        194 LAKPTPKAETTTKGPALTTPKEPTPTTPKEPASTTP---------------------- 229

Qy        723 ALHASGNEVPRGECSGPATVNNSSDTESIPSPHTEAAKDTGQNGPKPPATLGADGPPPGP 782
                  : |      | |: ::  |   |:| |  :   |    | |  |      | |
Db        230 -------KEPT-----PTTIKSAPTTPKEPAPTTTKSAPTTPKEPAPTTT-----KEPAP 272

Qy        783 PTPPRRTSRAPIEPTPASEATGAPTPP--PAPPSPSAPPPVVPKEEKEEETAAAPPVEEG 840
               ||          || | :    ||| |  ||| :|  | |  |||     |    |
Db        273 TTPKEPAPTTTKEPAPTT-TKSAPTTPKEPAPTTPKKPAPTTPKE--PAPTTPKEPTPTT 329

Qy        841 EEQKPPAAEELAVDTGKAEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGS 900
               ::  |  :| |  | |   |         | : ||     | |  | :       ||
Db        330 PKEPAPTTKEPAPTTPKEPAPT-------APKKPAPTTPKEPAPTTPKEPAPTTTKE--- 379

Qy        901 GRATTAKSSGAPQDSDSSATCSADEVDEAEGGDKNRLLSPRPSLLTPTGDPRANASPQKP 960
                : |     ||  : |: | : :                        ||    |  :|::|
Db        380 -PSPTTPKEPAPTTTKSAPTTTKEP--------------------APTTTKSAPTTPKEP 418

Qy        961 LDLKQLKQRAAAIPPIQVTKVHEP----PREDA--APTKPAPPAPPPPQNLQPESDAPQQ 1014
                                   |   ||    |:| |   | ||||  |. |    |:  ||
Db        419 ---------------SPTTTKEPAPTTPKEPAPTTPKKPAPTTPKEPAPTTPKEPAPTT 462

Qy       1015 PGSSPRGKSRSPAPPADKEAFAAEAQKLPGDPP----CWTSGLPFPVPPREVIKASPHAP 1070
                    : |||    ||      :||     |       |  | | | |:   :| |
Db        463 TKKPAPTAPKEPAPTTPKETAPTTPKKLTPTTPEKLAPTTPEKPAPTTPEELAPTTPEEP 522

Qy       1071 DPSAFSYAPPGHPLPLGLHDTARPVLPRPPTISNP-PPLISSAKHPSVLERQIGAISQGM 1129
                |:      |  | |      | |  |: |  : |  |  :: | |:
Db        523 TPTT-----PEEPAPT-TPKAAAPNTPKEPAPTTPKEPAPTTPKEPA------------ 563

Qy       1130 SVQLHVPYSEHAKAP-----VGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPESL 1184
                  | :    ||       | |:  | |  ||| ||      ::|:|      |  :
Db        564 ------PTTPKETAPTTPKGTAPTTLKEPAPTTPKKPAP------KELAPTTTKEPTSTT 611

Qy       1185 G---VPTAQEASVLRGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHGTPADVLYKG 1241
                  ||       :|||            |:|  | : | |      |   ||    ||
Db        612 SDKPAPTTP-----KGTA------------PTT--PKEPAPT-----TPKEPAPTTPKG 646

Qy       1242 TITRIIGEDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPP 1301
                |    : | :|:   :     |                : :|  | |       : :   |
Db        647 TAPTTLKEPAPTTPKKPAPKELAP-----------TTTKGPTSTT-----SDKPAPTTP 689

Qy       1302 HETA--APKRTYDMMEGRVGRAISSASIEGLMGRAIPP---ERHSPHHLKEQHHIRGSIT 1356
                 |||  ||          :                  ||   |  :|   ||   |  |
Db        690 KETAPTTPKEPAPTTPKKPAPTTPETP---------PPTTSEVSTPTTTKEPTTIHKSPD 740

Qy       1357 QGIPRSYVEAQEDYLRREAKLLKREGTP-----PPPPPSRDLTEAYKTQALGPLKLKPAH 1411
                :  |                    |  | ||     |  |    |:         |   ||
Db        741 ESTPE----------------LSAEPTPKALENSPKEPGVPTTKT-------PAATKPE- 776

Qy       1412 EGLVATVKEAGRSIHEIPREELRHTPE-LPLAPRPLKEGSIT------------------ 1452
                  :  | |:      :    :|| |||    ||:  || : |
Db        777 --MTTTAKD------KTTERDLRTTPETTTAAPKMTKETATTTEKTTESKITATTTQVTS 828

Qy       1453 ----QGTPLKYDTGAS-------TTGSKKHDVRSLIGSPGRTFPPVHPLDVMADARALER 1501
                    || |  |  :       || |      ::  |  |                |:  :|
Db        829 TTTQDTTPFKITTLKTTTLAPKVTTTKKTITTTEIMNKPEET-----------AKPKDR 876

Qy       1502 ACYEESLKSRPGTASSSGGS-IARGAPVIVPELGKPRQSPLTYEDHGAPFAGHLPRGSPV 1560
                |   ::   :|   : :     :   |  :| : ||: :|                ||
Db        877 ATNSKATTPKPQKPTKAPKKPTSTKKPKTMPRVRKPKTTP-------------TPRKMTS 923

Qy       1561 TMRE--PTPRLQEGSLSSS---KASQDRKLT-------------STPREIAKSPHSTVP 1601
                || |  || |: |  | ::     : : ||                 ||  : : ||  :|
Db        924 TMPELNPTSRIAEAMLQTTTRPNQTPNSKLVEVNPKSEDAGGAEGETPHMLLR-PHVFMP 982

Qy       1602 EHHPH----PISPYEHLLRGVSGVDLYRSHIPLAFDPTSIPRGIPLD 1644
                |  |      |  | : ::           |:  | |:|  | |:|
Db        983 EVTPDMDYLPRVPNQGIIIN-----------PMLSDETNICNGKPVD 1018

RESULT 49
US-10-084-846A-7
; Sequence 7, Application US/10084846A
; Publication No. US20040006026A1
; GENERAL INFORMATION:
;  APPLICANT: WEITNAUER, GABRIELE
;  APPLICANT:  MUHLENWEG, AGNES
;  APPLICANT:  TREFZER, AXEL
;  APPLICANT:  BECHTHOLD, ANDREAS
;  TITLE OF INVENTION: AVILAMYCIN DERIVATIVES
;  FILE REFERENCE: 1974-005
;  CURRENT APPLICATION NUMBER: US/10/084,846A
;  CURRENT FILING DATE:  2003-02-25
;  PRIOR APPLICATION NUMBER: PCT/EP01/09815
;  PRIOR FILING DATE: 2001-08-24
;  PRIOR APPLICATION NUMBER: DE 101 09 166.4
;  PRIOR FILING DATE: 2001-02-25
;  NUMBER OF SEQ ID NOS: 120
;  SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 7
;   LENGTH: 19652
```

```
;   TYPE: PRT
;   ORGANISM: Streptomyces viridochromogenes
;   FEATURE:
;   OTHER INFORMATION: Protein 2: amino acid sequence encoded by coding strand 2.
;   OTHER INFORMATION: Start codon: gat, Start position: nucleotide 2.
US-10-084-846A-7

  Query Match             3.1%;  Score 409;  DB 15;  Length 19652;
  Best Local Similarity   22.0%;  Pred. No. 4.3e-09;
  Matches  468;  Conservative  165;  Mismatches  808;  Indels  690;  Gaps  115;

Qy        708 SGVSGNEEEMVEE---------AEALH-----ASGNEVPRGECSGPATVNNSSDTESIPS 753
Db       1660 AAVTGHDSAGVEEPARPSGGRMSTALEWFKSSYSGSE--GGQC-----------VEVALC 1706

Qy        754 PHT---EAAKDTGQNGPK-----------PPATLGADGPPPGPPT----PPR-------- 787
Db       1707 PHTIHIRDSKNTPEDGPTLQVSPTAWRAFTSATTEARRPELTPPTVPVPPPRCGGPGRGP 1766

Qy        788 RTSRAPIEPTPASEATGAPTPPPAPPSPSAPPPVVPKEEKEEETAAAP--PVEEGEEQKP 845
Db       1767 RCRRRPVS-TRRPRCGGRKRRTGGTCSPASAPGVRPPLPCRAGGARGPSGPGPCGGHGRP 1825

Qy        846 PAAEELAVDTGKAEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGSGRAT- 904
Db       1826 PGVRRSAAAARRGSPRRRPAC------GPLRLPRSARAVRRDRARRRSRRRAAVGGRAAR 1879

Qy        905 ---TAKSSGAPQDSDSSATCSADEVDEAEGGDKNRLLSPRPSLLTPTGDPRANASPQKPL 961
Db       1880 CPRSARTHARPA---VPAPCAAPSAPSAPVCPVRRRSPPGPA--CPAWLPFPPPGPAHPL 1934

Qy        962 DLKQLK---QRAAAIPPIQVTKVHEPPREDAAPTKPAPPAPPPPQNLQPESDAPQQPGS- 1017
Db       1935 WLAPLRTPARPAGLLPPL------------ARPRPAPRVP----------LPVRPGTR 1970

Qy       1018 -SPRGKSRSPAPPADKEAFAAEAQKLPG--DPPCWTSGLPFPVPPREVIKASPHAPDPSA 1074
Db       1971 RTPRPSARSRRPPQRSRAAGAPSRTAPDARTPPCASPSCR---------RARPRGRPPRP 2021

Qy       1075 FSYAPP-----------GHPLPLGLHDTARPVL------PRPPTISNPPPLISSAKHPSV 1117
Db       2022 PCWAPPTRPCAGAPGVRGRPAGCGADRPVRPPCGTPWSPPRRRARRWPPPARAGGSAPG- 2080

Qy       1118 LERQIGAISQGMSVQLHVPYSEHAKAPVGPVTMGLPLPMDPKKLAPFSGVKQEQL----S 1173
Db       2081 ----CGTPRRGSPWAYRSGTARGRSAP------------PTRRRPWSGGSGDSSRSGPR 2123

Qy       1174 PRGQAGP--PESLGVPTAQ--------------------EASVLRGTALGSVPGGSITK 1210
Db       2124 PRRACPPRRPRAPGSPAARRRWPRRRPAPSSRRTRPAGRGRAAVSRAAHCASRSPGSSSA 2183

Qy       1211 GIPSTRVPSDSAITYRGSITHGTPADVLYKGTITRIIGEDSPSRLDRGREDSLPKGHVIY 1270
Db       2184 RWPD-RPPSGSAPRRLGARVRRWP------GSSCR-----APSR----RRGSSGAG---- 2223

Qy       1271 EGKKGHVLSYEGGMSVT---QCSKEDGRSSSGPPHETAAPKRTYDMMEG----RVGRAIS 1323
Db       2224 RGRSG--CRRPGSRAVRPGGACRRAPG-SPARPPRPAGTPRPPRGRGRGPRARRAGRSTR 2280

Qy       1324 SASIEGLMGRAIPPERHSPHHLKEQHHIRGSITQGIPRSYVEAQEDYLRREAKLLKREGT 1383
Db       2281 RSAARSRPGR--PP--RSP-----------GGAAGRPRPGRAAVAHRRGRPARSVR--GS 2323

Qy       1384 PPPPP---PSRDLTEAYKTQALGPLKLKPAHEGLVATVKEAGRSIHEIPREELRHTPELP 1440
Db       2324 PPRPPAPVPVRPAPRTACAAGLPPAPPRPA----------AGRASSSAPRPPRR--PGVP 2371

Qy       1441 LAPRPLKEGSITQGTPLKYDTGASTTGSKKHDVRSLIGSPGRTFPPVHPLDVMADARALE 1500
Db       2372 RGPPP----------PAR---------AARRPVR---WSGARACP----------RRCVR 2399

Qy       1501 RACYEESLKS------------RPGTASSSGGSIARGAPVIVPELGKP--RQSPLTYE 1544
Db       2400 RCCPVQPRRSAGRRPPSGRRSAARPRGAGAAGTS-RRRAP------GRPSGTRPSP---- 2448

Qy       1545 DHGAPFAGHLPRGSPVTMREPTPRLQEGSLSSSKASQDRKLTSTPREIAKSPHSTVPEHH 1604
Db       2449 ---PPPGAACPRGP--TAGPPGPPARAGPVAHG---------SVPGPPVR-PRRRVRDRR 2493

Qy       1605 PHPISPYEHLLRGVSGVDLYRSHIPLAFDPTSIPRGIPLDAAAAYYLPRHLAPNPTYPHL 1664
Db       2494 PPPAPPREAARPGPP---------PPAHSAPSRPR-----------PSHRVPASCHP-- 2530

Qy       1665 YPPYLIRGY----PDTAALENRQTIINDYITSQQMHHNTATAMAQR--ADMLRGLSPRES 1718
Db       2531 -GGGRIDGWCAGCPATLTIGRRANTCPLYVGSR-----GALGQADRLKAEIGRLNGPVRS 2584

Qy       1719 -----------SLALNYAAGP----------RGIIDLSQVPHLPVLVPPTPGTPA----- 1752
Db       2585 WPAPAVPDTRRSRAASPTTGPGRAGRDRGPGRGARS----------VPSRPPVPAWRRCR 2634

Qy       1753 TAMDRLAYLP---------TAPQPFS-----SRHSSSPLSPGGPTHLTKPTTTSSSERER 1798
Db       2635 ALLPRRSALPPWKRGGADSAPPNLFCAAGGLSRRSPSPWS-------ASPTSCRCAPDRG 2687

Qy       1799 DRDRERDRDREREKSILTSTTTVEHAPIWRPGT----------------EQSSGSSGSSG 1842
Db       2688 DRARQ---PRRRP----TAAAAGSRPPPARPGSAPPPSPGRRAGSRRRRRRSRGSGSPSG 2740

Qy       1843 GG----------------GGSSSRPASHSHAHQHSPISPRTQDALQQRPSVLHNTGMKGI 1886
Db       2741 TGRCPCWPRPVPSRTARCGRNPARRSAPCRSRRPRPRAPRT------------------- 2781

Qy       1887 ITAVEPSKPTVLRST--------STSSPVRPAATFPPATHCPLGGTLDGVYPTLMEPVLL 1938
Db       2782 ----GPPGRTSFRCTPRSPWSQGSASASVRREAPGRRARHLPRG---DGQRP-------- 2826

Qy       1939 PKEAPRVARPERPRADTGHAFLAKPPARSGLEPAS-SPSKGSEPRPLVPPVSGHAT---- 1993
Db       2827 ------ADQPRRGDDPSGTRALAE-PVRPGVAPGGRQASAGAGVRGGVPPGGGDTLRGTG 2879

Qy       1994 ---IARTPAKNLAPHHA-------SPDPPAPPASASDPHREKTQSKPFSIQELELRSLGY 2043
Db       2880 TCGVVRSSLGRLALRAARWALSSNSRRPPGGRAARQTGIRQ-------GLGRPRLRVL-- 2930

Qy       2044 HGSSYSPEGVEPVSPVSSPSLTHDKGLPKH--LEELDKSHLEG-------ELRPKQPGPV 2094
Db       2931 -GDCLPAQRVGP-------------QPGHRLVQRADGGPAEEFAGGPGVDLDPVVPGRA 2975

Qy       2095 K--------LGGEAAHL-------PHLRPLPESQPSSSPLLQTAPGVKGHQRVVTLAQHI 2139
Db       2976 RGGGGKLPDLGGHAAHLDEQIDEPPHRHGLAASDGDDQPRNTGAGG--GQHRVRDVA-HV 3032

Qy       2140 SEVIT-------QDYTRHHP-QQLSAPLPAPLYSFPGASCPVLDLR--RP-----PSDLY 2184
Db       3033 DVVAVGAQLGQLHDRREHVPGDRLGEP--------PGHPAEQADRRGARPGHVGDPQHAR 3084

Qy       2185 LPPP-----DHGAPARGSPHSEGGKRSPE---PNK-----TSVLGGGEDGIEPV----SP 2227
Db       3085 LHDALRGGLQHGLARREFGLAVGGRRPGRRRGPGRFTGRGRAVLQHAE--VDDVGTARGP 3142

Qy       2228 PEGMTEPGHSRSAV-------YPLLYRDG-----------------------EQTEPSR 2256
Db       3143 PQRVQQAAGRHRAVGLHVPFGDPPVAADGVDHRVRADGVDHVGQFPGRRGGDVQDEEPAR 3202

Qy       2257 MGSKSPGNTSQPPAFFSKLTESNSAMVKSKKQEINKKLNTHNRNEPEYNISQPGTEIFNM 2316
Db       3203 Q-----GEVRYPP--------GDRRHLGSGRGQVRAQLTAH---EP-----RPTSHENAL 3241

Qy       2317 PAITGTGLMTYRSQAVQEHASTNMGLEAIIRKALMGKYDQWEESPPLSANAFNPLNASAS 2376
Db       3242 ALIGHLLLLTSRCRAAVPGPAPDLPIGPVRHRRC---------APPTRRRRPGRSTRGCR 3292

Qy       2377 LPAAMPITAADGRSDHTLTSPGGGGKAKVSGRPSSRKA-KSPAPGLASGDRPPSVSSVHS 2435
Db       3293 RTGRRPAGRARAPRRRVPAEPWGGSRA--AGRHGRRSGPRSPTACTRSGRR---ARSRRT 3347
```

```
Qy       2436 EGDCNRRTPLTNRVWEDRPSSAGSTPFP--YNPLIMRLQAGVMASP------PPPG---- 2483
Db       3348 RVPCGRRAHTAAR----RPPRRLSPRFPSAASPRPHRGRARGTCSPRRADTRPRPAGCRR 3403

Qy       2484 -LPAGSG----PLAGPHHAWDE-----EPKP 2504
Db       3404 PAPAGTGRCREAVAGTRTRWGSAVHRCRPRP 3434


RESULT 50
US-10-124-557-74
; Sequence 74, Application US/10124557
; Publication No. US20020137894A1
;   GENERAL INFORMATION:
;        APPLICANT: Turner, Katherine
;                   Clark, Stephen C.
;                   Jacobs, Kenneth
;                   Hewick, Rodney M.
;                   Gesner, Thomas G.
;        TITLE OF INVENTION: Megakaryocyte Stimulating Factors
;        NUMBER OF SEQUENCES: 143
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Genetics Institute, Inc.
;             STREET: 87 CambridgePark Drive
;             CITY: Cambridge
;             STATE: Massachusetts
;             COUNTRY: U.S.A.
;             ZIP: 02140
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/124,557
;             FILING DATE: 16-Apr-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US 07/643,502
;             FILING DATE: 18-JAN-1991
;             APPLICATION NUMBER: US 07/546,114
;             FILING DATE: 29-JUN-1990
;             APPLICATION NUMBER: US 07/457,196
;             FILING DATE: 29-DEC-1989
;             APPLICATION NUMBER: US 07/390,901
;             FILING DATE: 08-AUG-1989
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Cserr, Luann
;             REGISTRATION NUMBER: 31,822
;             REFERENCE/DOCKET NUMBER: GI 5190
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (617)876-1170
;             TELEFAX: (617)876-5851
;   INFORMATION FOR SEQ ID NO: 74:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1038 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 74:
US-10-124-557-74

  Query Match             3.1%;  Score 408.5;  DB 13;  Length 1038;
  Best Local Similarity   21.0%;  Pred. No. 1.4e-10;
  Matches  272;  Conservative  132;  Mismatches  497;  Indels  395;  Gaps   58;

Qy        515 SSQE--------EKDEKEKEKEAEKEEEK------------PEVENDKEDLLKEK----- 549
Db         23 SSQELSCKGRCFESFERGRECDCDAQCKKYDKCCPDYESFCAEVKDNKKNRTKKKPTPKP 82

Qy        550 --TDDT-SGEDNDEKEAVASKGRKTANSQGRRKGRITRSMANEANSEEAITP-QQSAELA 605
Db         83 PVVDEAGSGLDNGDFKVTTPDTSTTQHNKVSTSPKIT--TAKPINPRPSLPPNSDTSKET 140

Qy        606 SMELNESSRWTEEEMETAKKGLLEHGRNWSAIARMVGSKTVSQCKNFYFNYKKRQNLDEI 665
Db        141 SLTVNKETTVETKETTTTNKQTSTDGKEKTTSAKETQSIEKTSAKDL-------APTSKV 193

Qy        666 LQQHKLKME---KERNARRKKKKAPAAASEEAAFPPVVEDEEMEASGVSGNEEEMVEEAE 722
Db        194 LAKPTPKAETTTKGPALTTPKEPTPTTPKEPASTTP------------------------- 229

Qy        723 ALHASGNEVPRGECSGPATVNNSSDTESIPSPHTEAAKDTGQNGPKPPATLGADGPPPGP 782
Db        230 -------KEPT-----PTTIKSAPTTPKEPAPTTTKSAPTTPKEPAPTTT-----KEPAP 272

Qy        783 PTPPRRTSRAPIEPTPASEATGAPTPP--PAPPSPSAPPPVVPKEEKEEETAAAPPVEEG 840
Db        273 TTPKEPAPTTTKEPAPTT-TKSAPTTPKEPAPTTPKKPAPTTPKE--PAPTTPKEPTPTT 329

Qy        841 EEQKPPAAEELAVDTGKAEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGS 900
Db        330 PKEPAPTTKEPAPTTPKEPAPT-------APKKPAPTTPKEPAPTTPKEPAPTTTKE--- 379

Qy        901 GRATTAKSSGAPQDSDSSATCSADEVDEAEGGDKNRLLSPRPSLLTPTGDPRANASPQKP 960
Db        380 -PSPTTPKEPAPTTTKSAPTTTKEP-------------------APTTTKSAPTTPKEP 418

Qy        961 LDLKQLKQRAAAIPPIQVTKVHEP----PREDA--APTKPAPPAPPPPQNLQPESDAPQQ 1014
Db        419 ---------------SPTTTKEPAPTTPKEPAPTTPKKPAPTTPKEPAPTTPKEPAPTT 462

Qy       1015 PGSSPRGKSRSPAPPADKEAFAAEAQKLPGDPP----CWTSGLPFPVPPREVIKASPHAP 1070
Db        463 TKKPAPTAPKEPAPTTPKETAPTTPKKLTPTTPEKLAPTTPEKPAPTTPEELAPTTPEEP 522

Qy       1071 DPSAFSYAPPGHPLPLGLHDTARPVLPRPPTISNP-PPLISSAKHPSVLERQIGAISQGM 1129
Db        523 TPTT-----PEEPAPT-TPKAAAPNTPKEPAPTTPKEPAPTTPKEPA------------ 563

Qy       1130 SVQLHVPYSEHAKAP-----VGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPESL 1184
Db        564 ------PTTPKETAPTTPKGTAPTTLKEPAPTTPKKPAP------KELAPTTTKEPTSTT 611

Qy       1185 G---VPTAQEASVLRGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHGTPADVLYKG 1241
Db        612 SDKPAPTTP-----KGTA------------PTT--PKEPAPT-----TPKEPAPTTPKG 646

Qy       1242 TITRIIGEDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPP 1301
Db        647 TAPTTLKEPAPTTPKKPAPKELAP-----------TTTKGPTSTT-----SDKPAPTTP 689

Qy       1302 HETA--APKRTYDMMEGRVGRAISSASIEGLMGRAIPP---ERHSPHHLKEQHHIRGSIT 1356
Db        690 KETAPTTPKEPAPTTPKKPAPTTPETP---------PPTTSEVSTPTTTKEPTTIHKSPD 740

Qy       1357 QGIPRSYVEAQEDYLRREAKLLKREGTP-----PPPPPSRDLTEAYKTQALGPLKLKPAH 1411
Db        741 ESTPE---------------LSAEPTPKALENSPKEPGVPTTKT-------PAATKPE- 776

Qy       1412 EGLVATVKEAGRSIHEIPREELRHTPE-LPLAPRPLKEGSIT------------------ 1452
Db        777 --MTTTAKD------KTTERDLRTTPETTTAAPKMTKETATTTEKTTESKITATTTQVTS 828

Qy       1453 ----QGTPLKYDTGAS-------TTGSKKHDVRSLIGSPGRTFPPVHPLDVMADARALER 1501
Db        829 TTTQDTTPFKITTLKTTTLAPKVTTTKKTITTTEIMNKPEET------------AKPKDR 876

Qy       1502 ACYEESLKSRPGTASSSGGS-IARGAPVIVPELGKPRQSPLTYEDHGAPFAGHLPRGSPV 1560
Db        877 ATNSKATTPKPQKPTKAPKKPTSTKKPKTMPRVRKPKTTP-------------TPRKMTS 923

Qy       1561 TMRE--PTPRLQEGSLSSSKASQDRKLTSTPREIAKSPHSTVPEHHPHPISPYEHLLRGV 1618
Db        924 TMPELNPTSRIAEAMLQT---------TTRPNQ---TPNSKLVE-----VNPKSEDAGGA 966
```

```
Qy       1619 SGVD---LYRSHIPLAFDPTSIPRGIPLDAAAAYYLPRHLAPNPTYPHLYPPYLIRGYPD 1675
Db        967 EGETPHMLLRPHV---FMPEVTP---DMD-----YLPR---------------------- 993

Qy       1676 TAALENRQTIINDYITSQ----QMHHNTATAMAQRA 1707
Db        994 ---VPNQGIIINPMLSDQYYNIDVPSRTARAITTRS 1026


RESULT 51
US-10-124-557-42
; Sequence 42, Application US/10124557
; Publication No. US20020137894A1
;  GENERAL INFORMATION:
;       APPLICANT: Turner, Katherine
;                  Clark, Stephen C.
;                  Jacobs, Kenneth
;                  Hewick, Rodney M.
;                  Gesner, Thomas G.
;       TITLE OF INVENTION: Megakaryocyte Stimulating Factors
;       NUMBER OF SEQUENCES: 143
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Genetics Institute, Inc.
;            STREET: 87 CambridgePark Drive
;            CITY: Cambridge
;            STATE: Massachusetts
;            COUNTRY: U.S.A.
;            ZIP: 02140
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/124,557
;            FILING DATE: 16-Apr-2002
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US 07/643,502
;            FILING DATE: 18-JAN-1991
;            APPLICATION NUMBER: US 07/546,114
;            FILING DATE: 29-JUN-1990
;            APPLICATION NUMBER: US 07/457,196
;            FILING DATE: 29-DEC-1989
;            APPLICATION NUMBER: US 07/390,901
;            FILING DATE: 08-AUG-1989
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Cserr, Luann
;            REGISTRATION NUMBER: 31,822
;            REFERENCE/DOCKET NUMBER: GI 5190
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (617)876-1170
;            TELEFAX: (617)876-5851
;  INFORMATION FOR SEQ ID NO: 42:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 1311 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 42:
US-10-124-557-42

  Query Match             3.1%;  Score 408.5;  DB 13;  Length 1311;
  Best Local Similarity   21.1%;  Pred. No. 1.8e-10;
  Matches  257;  Conservative  119;  Mismatches  489;  Indels  351;  Gaps   52;

Qy        518 EEKDEKEKEKEKEAEKEEEKPEVENDKEDLLKEKTDDTSGEDNDEKEAVASKGRKTANSQGR 577
Db        106 EVKDNK-KNRTKKKPTPKPPVVDEA----------GSGLDNGDFKVTTPDTSTTQHNKVS 154

Qy        578 RKGRITRSMANEANSEEAITP-QQSAELASMELNESSRWTEEEMETAKKGLLEHGRNWSA 636

Db        155 TSPKIT--TAKPINPRPSLPPNSDTSKETSLTVNKETTVETKETTTTNKQTSTDGKEKTT 212

Qy        637 IARMVGSKTVSQCKNFYFNYKKRQNLDEILQQHKLKME---KERNARRKKKKAPAAASEE 693
Db        213 SAKETQSIEKTSAKDL-------APTSKVLAKPTPKAETTTKGPALTTPKEPTPTTPKEP 265

Qy        694 AAFPPVVEDEEMEASGVSGNEEEMVEEAEALHASGNEVPRGECSGPATVNNSSDTESIPS 753
Db        266 ASTTP--------------------------------KEPT-----PTTIKSAPTTPKEPA 289

Qy        754 PHTEAAKDTGQNGPKPPATLGADGPPPGPPTPPRRTSRAPIEPTPASEATGAPTPP--PA 811
Db        290 PTTTKSAPTTPKEPAPTTT-----KEPAPTTPKEPAPTTTKEPAPTT-TKSAPTTPKEPA 343

Qy        812 PPSPSAPPPVVPKEEKEEETAAAPPVEEGEEQKPPAAEELAVDTGKAEEPVKSECTEEAE 871
Db        344 PTTPKKPAPTTPKE--PAPTTPKEPTPTTPKEPAPTTKEPAPTTPKEPAPT-------AP 394

Qy        872 EGPAKGKDAEAAEATAEGALKAEKKEGGSGRATTAKSSGAPQDSDSSATCSADEVDEAEG 931
Db        395 KKPAPTTPKEPAPTTPKEPAPTTTKE----PSPTTPKEPAPTTTKSAPTTTKEP------ 444

Qy        932 GDKNRLLSPRPSLLTPTGDPRANASPQKPLDLKQLKQRAAAIPPIQVTKVHEP----PRE 987
Db        445 --------------APTTTKSAPTTPKEP---------------SPTTTKEPAPTTPKE 474

Qy        988 DA--APTKPAPPAPPPPQNLQPESDAPQQPGSSPRGKSRSPAPPADKEAFAAEAQKLPGD 1045
Db        475 PAPTTPKKPAPTTPKEPAPTTPKEPAPTTTKKPAPTAPKEPAPTTPKETAPTTPKKLTPT 534

Qy       1046 PP----CWTSGLPFPVPPREVIKASPHAPDPSAFSYAPPGHPLPLGLHDTARPVLPRPPT 1101
Db        535 TPEKLAPTTPEKPAPTTPEELAPTTPEEPTPTT-----PEEPAPT-TPKAAAPNTPKEPA 588

Qy       1102 ISNP-PPLISSAKHPSVLERQIGAISQGMSVQLHVPYSEHAKAP-----VGPVTMGLPLP 1155
Db        589 PTTPKEPAPTTPKEPA------------------PTTPKETAPTTPKGTAPTTLKEPAP 629

Qy       1156 MDPKKLAPFSGVKQEQLSPRGQAGPPESLG---VPTAQEASVLRGTALGSVPGGSITKGI 1212
Db        630 TTPKKPAP------KELAPTTTKEPTSTTSDKPAPTTP-----KGTA------------- 665

Qy       1213 PSTRVPSDSAITYRGSITHGTPADVLYKGTITRIIGEDSPSRLDRGREDSLPKGHVIYEG 1272
Db        666 PTT--PKEPAPT-----TPKEPAPTTPKGTAPTTLKEPAPTTPKKPAPKELAP------- 711

Qy       1273 KKGHVLSYEGGMSVTQCSKEDGRSSSGPPHETA--APKRTYDMMEGRVGRAISSASIEGL 1330
Db        712 -----TTTKGPTSTT-----SDKPAPTTPKETAPTTPKEPAPTTPKKPAPTTPETP---- 757

Qy       1331 MGRAIPP---ERHSPHHLKEQHHIRGSITQGIPRSYVEAQEDYLRREAKLLKREGTP--- 1384
Db        758 -----PPTTSEVSTPTTTKEPTTIHKSPDESTPE---------------LSAEPTPKAL 796

Qy       1385 --PPPPPSRDLTEAYKTQALGPLKLKPAHEGLVATVKEAGRSIHEIPREELRHTPE-LPL 1441
Db        797 ENSPKEPGVPTTKT-------PAATKPE---MTTTAKD------KTTERDLRTTPETTTA 840

Qy       1442 APRPLKEGSIT----------------------QGTPLKYDTGAS-------TTGSKKHD 1472
Db        841 APKMTKETATTTEKTTESKITATTTQVTSTTTQDTTPFKITTLKTTTLAPKVTTTKKTIT 900

Qy       1473 VRSLIGSPGRTFPPPVHPLDVMADARALERACYEESLKSRPGTASSSGGS-IARGAPVIVP 1531
Db        901 TTEIMNKPEET------------AKPKDRATNSKATTPKPQKPTKAPKKPTSTKKPKTMP 948

Qy       1532 ELGKPRQSPLTYEDHGAPFAGHLPRGSPVTMRE--PTPRLQEGSLSSS---KASQDRKLT 1586
Db        949 RVRKPKTTP------------TPRKMTSTMPELNPTSRIAEAMLQTTTRPNQTPNSKLV 995

Qy       1587 --------------STPREIAKSPHSTVPEHHPH----PISPYEHLLRGVSGVDLYRSHI 1628
Db        996 EVNPKSEDAGGAEGETPHMLLR-PHVFMPEVTPDMDYLPRVPNQGIIIN----------- 1043
```

```
Qy       1629 PLAFDPTSIPRGIPLD 1644
              |:  | |:|  | |:|
Db       1044 PMLSDETNICNGKPVD 1059
```

RESULT 52
US-10-205-331-66
; Sequence 66, Application US/10205331
; Publication No. US20040058326A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; APPLICANT: Lee, Kevin
; APPLICANT: Dixon, Alistair
; APPLICANT: Brooksbank, Robert
; APPLICANT: Pinnock, Robert
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
; FILE REFERENCE: WL-A-018199
; CURRENT APPLICATION NUMBER: US/10/205,331
; CURRENT FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: GB 0118354.0
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 117
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 66
; LENGTH: 2364
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Microtubule associated protein IB
US-10-205-331-66

Query Match 3.1%; Score 406.5; DB 12; Length 2364;
Best Local Similarity 19.3%; Pred. No. 4.6e-10;
Matches 489; Conservative 326; Mismatches 995; Indels 729; Gaps 107;

```
Qy         68 LLSEFQPGNERSQEL----------------------HLRPESHSYLPEL-GKSEM---- 100
              |||   | |: |  |                      :::  | | |||: | ||
Db         79 LLSTTHPANKASLTLFCPEEGDWKNSNLDRHNLQDFINIKLNSASILPEMEGLSEFTEYL 138

Qy        101 -EFIESKRPRLELLPDP-------LLRP-----------SPLLATGQ-----PAGSED-- 134
               | :|   |  ::|  |        | :|           | | |          |||
Db        139 SESVEVPSP-FDILEPPTSGGFLKLSKPCCYIFPGGRGDSALFAVNGFNMLINGGSERKS 197

Qy        135 ----------------LTK--DRSLTG----------KLEPVSPPSPPHTD-------PE 159
                             ||   | :| |          :||  |  |  ::|       |:
Db        198 CFWKLIRHLDRVDSILLTHIGDDNLPGINSMLQRKIAELEEESQGSTSNSDWMKNLISPD 257

Qy        160 LEL----VPPRLSKEELIQNMDRVDREITMVEQQISKLKKKQQQLEEEAAKPPEP----E 211
              | :    ||  |   |    | |   |     | ::||  | : |         ||    :
Db        258 LGVVFLNVPENLKNPEPNIKMKRSTEEACFTLQYLNKLSMKPEPLFRSVGNAIEPVILFQ 317

Qy        212 K---------PVSPPPIESKHRSLVQIIYDENRKKAEAAHRILEGLGPQVELPLYNQPSD 262
              |          ::|     : :  :|     |: |||    ::   | :|::|
Db        318 KMGVGKLKMYVLNPVKSSKEMQYFMQQWTGTNKDKAE----LILPNGQEVDIP------- 366

Qy        263 TRQYHENIKINQAMRKKLILYFKRRNHARKQWKQKFCQRYDQ--LMEALEKKVERIENNP 320
                    |:      :: :   | | |  :  |      |  ::| || |:: ::
Db        367 ---------ISYLASVSSLIVWHPANPAEKIIRVLFPGNSTQYNILEGLE-KLKHLDFLK 416

Qy        321 RRRAKESKVREYYEKQFPEIRKQRELQERMQSRVGQRGSGLSMSAARSEHEVSEIIDGLS 380
              :  | :  :    :  | : || :|::|  ||   : :   :|:      | |     :
Db        417 QPLATQKDLTG--QVSTPPV-KQVKLKQRADSRESLKPATKPLSSKSVRKESKEEAPEAT 473

Qy        381 EQENLEKQMRQLAVIPPMLYDADQQRIKFINMNGLMADPMKVYKDR--QVMNMWSEQEKE 438
              :   :||          | :   ::  :|               ||:  :| :  |  |||
Db        474 KASQVEK--------TPKVESKEKVIVK--------------KDKPGKVESKPSVTEKE 510

Qy        439 TFREKFMQHPKNFGLIASFLERKTVAECVLYYYLTKKNENYKSLVRRSYRRRGKSQQQQQ 498
                  | | |     : : :  |   |       :|      :|::  : : : :::::
Db        511 V-PSKEEQSP-----VKAEVAEKAATE-------SKPKVTKDKVVKKEIKTKPEEKKEEK 557

Qy        499 QQQQQQQQQQQQPMPRSSQEEKDEKEKEKEAEKEEEKPEVENDKEDLLKEKTDDTSGEDN 558
```

```
Db        558 PKKEVAKKEDKTPL------KKDEKPKKEEAKKEIKKEIKKEEKKELKKEVKKETPLKD- 610

Qy        559 DEKEAVASKGRKTANSQGRRKGRITRSMANEANSEEAITPQQSAELASMELNESSRWTEE 618
                                    |  : :   |  ||   |::  :  | :: :|:  ::
Db        611 -------------------AKKEVKKDEKKEVKKEEK-EPKKEIKKISKDIKKSTPLSDT 650

Qy        619 EMETAKKGLLEHGRNWSAIARMVGSKTVSQCKNFYFNYKKRQNLDEILQQHKLKMEKERN 678
              :   | |           :|:                   |:     | :   ||| : :
Db        651 KKPAALK---------PKVAK----------------KEEPTKKEPIAAGKLKDKGKVK 684

Qy        679 ARRKKKKAPAAASEEAAFPPVVEDEEMEASGV--SGNEEEMVEEAEALHASGNEVPRGEC 736
              :|: |   ||:         |        |:|| ||  :|:                | |
Db        685 VIKKEGKTTEAAATAVGTAAVA-----AAAGVAASGPAKEL-----------EAERSLM 727

Qy        737 SGPATVNNSSDTESIPSPHTEAAKDTGQNGPKPPATLGADGPPPGPPTPPRRTSRAPIEP 796
              | |  :  : | | : :   : |||      ||   | |                  ::
Db        728 SSPEDL--TKDFEELKAEEIDVAKDI-----KPQLELIED-------------EEKLKE 766

Qy        797 TPASEATGAPTPPPAPPSPSAPPPVVPKEEKEEETAAAPPVE-----EGEEQKPPAAEEL 851
              |   ||                  |: || :  : :|  | |     ||| :     |||
Db        767 TEPGEAY----------------VIQKETEVSKGSAESPDEGITTTEGEGECEQTPEEL 809

Qy        852 AVDTGKAEEPVKSECTEEAE----EGPAKGKDAEAA--EATAEGALKAEKKEGGSGRATT 905
                    |||: :  :: |     ||     : :||   |  ||     | :| |   :
Db        810 --------EPVEKQGVDDIEKFEDEGAGFEESSEAGDYEEKAETEEAEEPEEDGEDNVSG 861

Qy        906 AKSSGAPQDSDSSATCSADEVDEAEGGDKNRLLSPRPSLLTPTGDPRANASPQKPLDLKQ 965
              : |  :| : :  |   ||              :   | |:   :|| ||     : |: :
Db        862 SASKHSPTEDEEIAKAEADV----------HIKEKRESV--ASGDDRAEEDMDEALEKGE 909

Qy        966 LKQRAAAIPPIQVTKVHEPPREDAAPTK------------PAPPAPPPPQNLQPESDAPQ 1013
              :| :      :  |  :   ||  | |              | |                :
Db        910 AEQ-SEEEGEEEEDKAEDAREEDHEPDKTEAEDYVMAVVDKAAEAGVTEDQYDFLGTPAK 968

Qy       1014 QPG-SSPRGKSRSPAPP---------ADKEAFAAEAQKLPGDPPCWTSGLPFPVPPREVI 1063
              ||| ||   || ||            :: || |:: :     |  :|:   :      |:
Db        969 QPGVQSP---SREPASSIHDETLPGGSESEATASDEENREDQPEEFTATSGYTQSTIEI- 1024

Qy       1064 KASPHAPDPSAFSYAPPGHPLPLGLHDTARPVLP----RPPTISNPPPLISSAKHPSVLE 1119
                                    | |:    | | |:       | |      ::  |: | |
Db       1025 ---------------SSEPTPMDEMSTPRDVMTDETNNEETESPSQEFVNITKYESSL- 1067

Qy       1120 RQIGAISQGMSVQLHVPYSEHAKAPVGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRG--Q 1177
                 ||  |  :   ::  :.        | |          ::| | :::::  |
Db       1068 -----YSQEYSKPVVASFNGLSDGSKTDATDGRDYNASASTISPPSSMEEDKFSKSALRD 1122

Qy       1178 AGPPESLGVPTAQEASV----------LRGTALGSVPGGSITK------GIPSTRVPSDS 1221
              |  ||   | |  |  :          :  :|   |  | |        :  :  |::
Db       1123 AYRPEETDVKTGAELDIKDVSDERLSPAKSPSLSPSPPSPIEKTPLGERSVNFSLTPNEI 1182

Qy       1222 AITYRGSITHGTPADVLYKGTITRIIGEDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYE 1281
                :  |  |       |:  |     :: |   |  ::  |   |   |   |   ||  |:
Db       1183 KASAEGEAT-----AVVSPGVTQAVVEEHCASPEEKTLEVVSPSQSV--TGSAGHTPYYQ 1235

Qy       1282 GGMSVTQCSKEDGRSSSGPPHET----AAPKRTYDMMEGRVGRAISSAS--------IEG 1329
                    |  | :||  |   |    | |  :::  | :      || |          |
Db       1236 --------SPTDEKSSHLPTEVTENAQAVPV-SFEFTEAKDENERSSISPMDEPVPDSES 1286

Qy       1330 LMGRAIPPERHSPHHLKEQHHIRGSITQGIPRSYVEAQEDYLRREAKLLKREGTPP---- 1385
              : : : | |  |        : ||           | ||:|  : | | |  |
Db       1287 PIEKVLSPLRSPP--------LIGS---------ESAYEDFLSADDKALGRRSESPFEGK 1329

Qy       1386 --------PPPPSRDLT-EAYKTQALGPLKLKPAHEGLVATVKEAGRSIHEIPREELRHT 1436
                     | ||| : |:       | :   | :   ::           | :
Db       1330 NGKQGFSDKESPVSDLTSDLYQD------KQEEKRAGFIPIKEDFSPEKKASDAEIMSSQ 1383

Qy       1437 PELPLAPRPL-KEGSITQGTPLKYDTGASTTGSKKHDVRSLIGSPGRTFPPVHPLDVMAD 1495
              | |  | |  :|| ||           |  || | |  :  |     :     |:|  |
Db       1384 SALALDERKLGGDGSPTQ-------VDVSQFGSFKEDTKMSISEGTVSDKSATPVDEGA- 1435

Qy       1496 ARALERACYEESLKSRPGTASSSGGSIARGAPVIVPELGKPRQSPLTYEDHGAPFAGHLP 1555
                    |::     | || |  |:|  :    ||      ||  : :  |:| :  :
```

```
Db       1436 ---------EDTYSHMEGVASVSTASVATSS---FPEPTTDDVSPSLHAEVGSPHSTEVD 1483

Qy       1556 RGSPVTMREPTPRLQEGSLSSSKASQDRKLTSTPREI----AKSPHSTVPEHHPHPISPY 1611
              |:: :     ||  :| ||    | :: :| :     |||  : | :|
Db       1484 DSLSVSVVQTPTTFQETEMSPSKEECPRPMSISPPDFSPKTAKS-RTPVQDHRSE----- 1537

Qy       1612 EHLLRGVSGVDLYRSHIPLAFDPTSIPRGIPLDAAAAYYLPRHLAPNPTYPHLYPPYLIR 1671
                          :| : : |   |    : :| :           :| :| :
Db       1538 ------------QSSMSIEFGQESPEHSLAMDFS---------RQSPDHPTV------- 1568

Qy       1672 GYPDTAALENRQTIINDYITSQQMHHNTATAMAQRADMLRGLSPRESSLALNYAAGPRGI 1731
              |      ||  | : ||                                    |  |
Db       1569 GAGMLHITENGPTEV-DY-----------------------------------SPSDI 1590

Qy       1732 IDLSQVPHLPVLVPPTPGTPATAMDRLAYLPTAPQPFSSRHSSSPLSPGGPTHLTKPTTT 1791
               | |    |  :|||     |  : |: | :  |  :|  :||  :|      :  |
Db       1591 QDSS----LSHKIPPTEEPSYTQDNDLSELISVSQVEASPSTSSAHTPSQIASPLQEDTL 1646

Qy       1792 SSSERERDRD------RERDRDRERE----KSILTSTTTVEHAPIWRPG-TEQSSGSSGS 1840
              |    ||          |: :  | |    || ::  |  | :| : || :: :||:  |
Db       1647 SDVVPPRDMSLYASLASEKVQSLEGEKLSPKSDISPLTPRESSPTYSPGFSDSTSGAKES 1706

Qy       1841 SGGGGGSSSRPASHSHAHQHSPISPRTQDALQQRPSVLHNTGMKGIIT--AVEP------ 1892
              :    ||| |   : |  :    |    | :|       |:  :   :|  :||
Db       1707 TAAYQTSSSPPIDAAAAEPYGFRSSMLFDTMQ------HHLALSRDLTTSSVEKDNGGKT 1760

Qy       1893 --------SKPTVLRSTSTSSPVRPAATFPPATHCPLGGTLDGVYPTLMEPVLLPKEAPR 1944
                      ||     ::| ||     :          : | |              :
Db       1761 PGDFNYAYQKP----ESTTESPDEEDYDYESHEKTIQAHDVGGYY-----------YEK 1804

Qy       1945 VARPERPRADTGHAF-----LAKPPARSGL------EPASSPSKG------SEPRPLVPP 1987
                |  :   |:|:::       | |    |       :   :| :|      ||     |
Db       1805 TERTIKSPCDSGYSYETIEKTTKTPEDGGYSCEITEKTTRTPEEGGYSYEISEKTTRTPE 1864

Qy       1988 VSGHATIARTPAKNLAPH------------HASPDPPAPPASASDPHREKTQSKP----F 2031
              |||:       :: |              |   |      | |    ||  | |    :
Db       1865 VSGYTYEKTERSRRLLDDISNGYDDTEDGGHTLGD-----CSYSYETTEKITSFPESESY 1919

Qy       2032 SIQELELRSLGYHGSSYSPEGVEPVSPVSSPS----------LTHDKGLPKH-------- 2073
              | :     :     |:|  | :| ::     |              | ::  |
Db       1920 SYETTTKTTRSPDTSAYCYETMEKITKTPQASTYSYETSDRCYTPERKSPSEARQDVDLC 1979

Qy       2074 -LEELDKSHLEGELRPK--QPGPVK-LGGEAAHLPHLRPLPESQPSSSPLLQTAPGVKGH 2129
               :   :  | : || |    | |::   ||       ||| :|  :  |   :    :|
Db       1980 LVSSCEFKHPKTELSPSFINPNPLEWFAGEEPTEESERPLTQSGGAPPP---SGGKQQGR 2036

Qy       2130 QRVVTLAQHISEVITQDYTRHHPQQLSAPLPAPLYSFPGASCPV-LDLRRP----PSDLY 2184
              |   |    :||          | |  : :|      |  :    ||         |:|
Db       2037 QCDETPPTSVSE--------SAPSQTDSDVPPETEECPSITADANLDSEDESETIPTDKT 2088

Qy       2185 L------PPPDHGAPARGSPHSEGGKRSPEPNKTSVLGGGEDGIEPVSPPEGMT------ 2232
              :      |||   || :        ||| |    |          :  || :
Db       2089 VTYKHMDPPP---APMQ--------DRSPSPRHPDV---------SMVDPEALAIEQNLG 2128

Qy       2233 --------------EPGHSRSAVYPLLYRDGEQTEPSRMGSKSPGNTSQPPAFFSKLTES 2278
                           :||    :  |:   ||:           |   |  |    : ::
Db       2129 KALKKDLKEKAKTKKPGTKTKSSSPVKKGDGK---------SKPSAASPKPGALKESSDK 2179

Qy       2279 NSAMVKSKKQE-INKKLNT 2296
                |  :   ||:| : | : |
Db       2180 VSRVASPKKKESVEKAMKT 2198
```

RESULT 53
US-10-196-935A-2
; Sequence 2, Application US/10196935A
; Publication No. US20030082720A1
; GENERAL INFORMATION:
; APPLICANT: Lifton, Richard P
; APPLICANT: Wilson, Frederick H
; APPLICANT: Choate, Keith
; APPLICANT: Ishikawa, Kazuhiko
; APPLICANT: Nelson-Williams, Carole

; TITLE OF INVENTION: COMPOSITIONS METHODS AND KITS RELATING TO TREATING AND DIAGNOSING
; TITLE OF INVENTION: HYPERTENSION
; FILE REFERENCE: 044574-5113
; CURRENT APPLICATION NUMBER: US/10/196,935A
; CURRENT FILING DATE: 2002-10-25
; PRIOR APPLICATION NUMBER: US 60/306,084
; PRIOR FILING DATE: 2001-07-17
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 2382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-196-935A-2

Query Match 3.1%; Score 406; DB 14; Length 2382;
Best Local Similarity 19.6%; Pred. No. 4.9e-10;
Matches 513; Conservative 291; Mismatches 943; Indels 872; Gaps 114;

```
Qy        147 PVSPPSPPHTDPELELVPPRLSKEELIQNMDRVDREITMVEQQISKLKKKQQQLEEEAAK 206
              |:| | |       :  ||   :|| :                |:::        |:|
Db         99 PLSLPQPSIPAAVPQSAPPEPHREETV---------TATATSQVAQQPPAAAAPGEQAVA 149

Qy        207 PPEP---------EKPVSPPPIESKHRSLVQIIYDENRKKAEAAHRILEGLGPQVELPLY 257
               | |          ::||| | :                          :| : | |
Db        150 GPAPSTVPSSTSKDRPVSQPSL---------------------------VGSKEEPPPA 181

Qy        258 NQPSDTRQYHENIKINQAMRKKLILYFKRRNHARKQWKQKFCQRYDQLMEALEKKVERIE 317
                |                                 : |: :::  |: | : | || |    :
Db        182 RSGSG-------------------------GGSAKEPQEERSQQQDDI-EELETKAVGMS 215

Qy        318 NNPRRRAKESKV-REYYEKQFPEIRKQR-------ELQERMQSRVGQRGSGLSMSAARSE 369
              |: |    : :: |  ::  :  :  :        |||:|            |: |  :
Db        216 NDGRFLKFDIEIGRGSFKTVYKGLDTETTVEVAWCELQDRK----------LTKSERQRF 265

Qy        370 HEVSEIIDGLSEQENLEKQMRQLAVIPPMLYDADQQRIK----FINMNGLM-ADPMKVYK 424
               | :|:: || :  |: :            ||: :  :|     : :  || :  :| |
Db        266 KEEAEMLKGL-QHPNIVR-----------FYDSWESTVKGKKCIVLVTELMTSGTLKTYL 313

Qy        425 DR------QVMNMWSEQEKETFREKFMQHPK----------------------NFGLIA 455
               |      :|:  |  |  :  :    : |                        : ||
Db        314 KRFKVMKIKVLRSWCRQILKGLQFLHTRTPPIIHRDLKCDNIFITGPTGSVKIGDLGL-- 371

Qy        456 SFLERKTVAECVL--------YYYLTKKNENY------KSLVRRSYRRRGKSQQQQQQQQ 501
              : |:| : |: |:           |  | :|:          ::  :       |: |  |
Db        372 ATLKRASFAKSVIGTPEFMAPEMYEEKYDESVDVYAFGMCMLEMATSEYPYSECQNAAQI 431

Qy        502 QQQQQQQQQP-------MPRSSQ-------EEKDEKEKEKEAEKEEEKPEVENDKEDLLK 547
               ::     :|      :|   :       : |||:   |:          |    : :| :
Db        432 YRRVTSGVKPASFDKVAIPEVKEIIEGCIRQNKDERYSIKDLLNHAFFQEETGVRVELAE 491

Qy        548 EKTDDTSGEDNDEKEAVASKGRKTANSQGRRKGRITRSMANEANSE-EAITPQQSAELAS 606
              |          |: | :| |          : ||:   : | | : : |   |:  |:
Db        492 E----------DDGEKIAIKLWLRIEDIKKLKGKYKDNEAIEFSFDLERDVPEDVAQ--- 538

Qy        607 MELNESSRWTEEEMETAKKGLLEHGRNWSAIARMVGSKTVSQCKNFYFNYKKRQNLDEIL 666
               |: ||    | : :|  | : :                  ||  |      :||:    :
Db        539 -EMVESGYVCEGDHKTMAKAIKDR--------------VSLIK------RKREQRQLVR 576

Qy        667 QQHKLKMEKERNARRKKKKA----------------PAAASEEAAFPPVVEDEEMEASG 709
               :: : | ::| : ::: :::                | |::  |:    || || ||
Db        577 EEQEKKKQEESSLKQQVEQSSASQTGIKQLPSASTGIPTASTTSASVSTQVEPEEPEA-- 634

Qy        710 VSGNEEEMVEEAEALHASGNEVPRGECSGPATVN--NSSDTESIPSPHTEAAKDTGQNGP 767
                 :::   ::       |    |  |:  |    | :  :|  | |  | | | |  ||
Db        635 -DQHQQLQYQQPSISVLSDGTVDSGQGSSVFTESRVSSQQTVSYGSQH-EQAHSTGTVPG 692

Qy        768 KPPATLGADGPPPG--PPT-----------------------------------PPR 787
               |:|: |   | |  ||:                                    ||:
Db        693 HIPSTVQAQSQPHGVYPPSSVAQGQSQGQPSSSSLTGVSSSQPIQHPQQQQGIQQTAPPQ 752

Qy        788 RTSRAPIEPT-PASEATGAPTPPPAPPSPSAPPPVVPKEEKEEETAAAPPVE--EGEEQK 844
              :| :  :  |  :|||| |      | |    | |:|:    ::   : ||   :|| |
```

```
Db        753 QTVQYSLSQTSTSSEATTA-----QPVSQPQAPQVLPQVSAGKQLPVSQPVPTIQGEPQI 807
Qy        845 PPAAEE--LAVDTGKAEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGSGR 902
Db        808 PVATQPSVVPVHSGAHFLPV----------------------------------GQPL 831

Qy        903 ATTAKSSGAPQDSDSSATCSADEVDEAEGGDKNRLLSPRPSLLTPTGDPRANASPQKPLD 962
Db        832 PTPL----LPQYPVSQIPISTPHVSTAQTG-----FSSLPITMA--------AGITQPLL 874

Qy        963 LKQLKQRAAAIP------PIQVTKVHEPPREDAAPTKPAPPAPPPPQNLQPESDAPQQPG 1016
Db        875 TLASSATTAAIPGVSTVVPSQLPTLLQPVTQ--------LPSQVHPQLLQP---AVQSMG 923

Qy       1017 SSPRGKSRSPAPPADKEAFAAEAQKLPGDPPCWTSGLPFPVPPREVIKASPHAPDPSAFS 1076
Db        924 I-----------PANL-GQAAEVPLSSGD--VLYQGFPPRLPPQY----------PGDSN 959

Qy       1077 YAPPGHPLPLGLHDTARPVL--PRPPTISNPPPLISSAKHPSV---LERQIGAISQGMSV 1131
Db        960 IAPSSNVASVCIHST---VLSPPMPTEVLATPGYFPTVVQPYVESNLLVPMGGV--GGQV 1014

Qy       1132 QLHVPYSEHAKAPVGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPESLGV----- 1186
Db        1015 QVSQPGGSLAQAPTTSSQQAV--------LESTQGV--------SQVAPAEPVAVAQPQA 1058

Qy       1187 --PTAQEASVLRGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHGTPADVLYKGTIT 1244
Db        1059 TQPTTLASSV--DSAHSDVASG-MSDG--NENVPSSSG--------------RHEGRTT 1098

Qy       1245 RIIGEDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKE------------ 1292
Db        1099 KRHYRKSVRSRSRHEKTSRPKLRILNVSNKGD--------RVVECQLETHNRKMVTFKFD 1150

Qy       1293 -DGRSSSGPPHETAAPKRTYDMMEGRVGRAISSASIEGLMGRAIPPERHSPHHLKEQHHI 1351
Db        1151 LDGDN----PEEIATIMVNNDFI-----LAIERESFVDQVREII---EKADEMLSEDVSV 1198

Qy       1352 RGSITQGIPRSYVEAQEDYLRREAKLLKREGTPPPPPPSRDLTEAYKTQALGPLKLKPAH 1411
Db        1199 EPEGDQGL--ESLQGKDDYGFSGSQKLEGEFKQPIPASSM-----------PQQIGIPT 1244

Qy       1412 EGLVATVKEAGRS--IHEIPREELRHTPELPLAPRPLKEGSITQGTPLKYDTGASTTGSK 1469
Db        1245 SSLTQVVHSAGRRFIVSPVPESRLRESKVFP--------SEIT-------DTVAASTAQS 1289

Qy       1470 -----KHDVRSL------------------------IGSPGRTFPPVHPLDVMADARALE 1500
Db        1290 PGMNLSHSASSLSLQQAFSELRRAQMTEGPNTAPPNFSHTGPTFPVVPP---------- 1338

Qy       1501 RACYEESLKSRPGTASSSGGSIARGAP-------VIVPELGKPRQSPL--TYEDHGAPFA 1551
Db        1339 ---FLSSIAGVPTTAAATAPVPATSSPPNDISTSVIQSEVTVPTEEGIAGVATSTGVVTS 1395

Qy       1552 GHLPRGSPVTMREPTPRLQEGSLSSSKASQDRKLTSTPREIA---KSPHSTVPEHHPHPI 1608
Db        1396 GGL----------PIPPVSESPVLSSVVSS----ITIPAVVSISTTSPSLQVPTSTSEIV 1441

Qy       1609 SPYEHLLRGVSGVDLYRSHIPLAFDPTSIPRGIPLDAAAAYYLPRHLAPNPTYPHLYPPY 1668
Db        1442 ---------VSSTALYPS--------------VTVSATSASAGGSTATPGPK-----PPA 1473

Qy       1669 LI--------------------RGYPDTAALENRQ-----------TIINDYITSQQMH 1696
Db        1474 VVSQQAAGSTTVGATLTSVSTTTSFPSTASQLSIQLSSSTSTPTLAETVVVSAHSLDKTS 1533

Qy       1697 HNTATAMA-----------------QRADMLRGLSPR------ESSLALNYAAGPRGIID 1733
Db        1534 HSSTTGLAFSLSAPSSSSSPGAGVSSYISQPGGLHPLVIPSVIASTPILPQAAGPTSTPL 1593

Qy       1734 LSQVPHLPVLVPPTPGTPATAMDRLAYLPTAPQPFSSRHSSSPLSPGGP-THLTKPTTTS 1792
Db        1594 LPQVPSIPPLVQPVANVPAV---QQTLIHSQPQP--------ALLPNQPHTHCP------ 1636

Qy       1793 SSERERDRDRERDRDREREKSILTSTTTVEHAPIWRPGTEQSSGSSGSSGGGGGSSSRPA 1852
Db        1637 ----------EVDSDTQPKAPGIDDIKTLEE-------KLRSLFSEHSSSGA-------- 1671

Qy       1853 SHSHAHQHSPISPRTQDALQQRPSVLHNTGMKGI-ITAVEPSKPTVLRSTSTSSPVRPAA 1911
Db        1672 ------QHASVSLETS-------LVIESTVTPGIPTTAVAPSK--LLTSTT--------S 1708

Qy       1912 TFPPATHCPLGGTLDGVYPTLMEPVLLPKEAPRVARPERPRADTGHAFLAKPPARSGLEP 1971
Db        1709 TCLPPTNLPLG--------------------TVALPVTPVVTPGQVSTPVSTTTSGVKP 1747

Qy       1972 ASSPSKGSEPRPLVPPVSGHATIARTPAKNLAPHHASPDPPAPPASASDPHREKTQS-KP 2030
Db        1748 GTAPSKPPLTKAPVLPVGTELPAGTLPSEQL--------PPFPGPSL-------TQSQQP 1792

Qy       2031 FSIQELELRSLGYHGSSYSPEGVEPVSPVSSPSLTHDKGLPKHLEELDKSHLEGELRPKQ 2090
Db        1793 LEDLDAQLR------RTLSPEMITVTSAV------------------------------ 1815

Qy       2091 PGPVKLGGEAAHLPHLRPLPE---SQPSSSPLLQTAPGV----KGHQRVVTLAQHISEVI 2143
Db        1816 -GPVSMAAPTA-ITEAGTQPQKGVSQVKEGPVLATSSGAGVFKMGRFQVSVAA------- 1866

Qy       2144 TQDYTRHHPQQLSAPLPAPLYSFPGASCPVLDLRRPPSDLYLPPP----------DHGAP 2193
Db        1867 --DGAQKEGKNKSEDAKSVHFESSTSESSVLSSSSPESTLVKPEPNGITIPGISSDVPES 1924

Qy       2194 ARGSPHSEGGKRSPEPNKTS--------------VLGGGEDGIE------PVSPPEGMTE 2233
Db        1925 AHKTTASEAKSDTGQPTKVGRFQVTTTANKVGRFSVSKTEDKITDTKKEGPVASPPFMDL 1984

Qy       2234 PGHSRSAVYPLLYRDGEQTEPSRMGSKSPGNTSQP-PAFFSKLTESNSAMVKSKKQEINK 2292
Db        1985 EQAVLPAVIPKKEKP-ELSEPSHLN----GPSSDPEAAFLSRDVDDGSGSPHSPHQLSSK 2039

Qy       2293 KLNTHNRNEPEYNISQPGTEIFNMPAITGTGLMTYRSQAVQ--------EHASTNMGLEA 2344
Db        2040 SL-------PSQNLSQSLSNSFNSSYMSSDNESDIEDEDLKLELRRLRDKHLKEIQDLQS 2092

Qy       2345 IIRKALMGKYDQWEESPPLSANAFNPLNASASLPAAMPITAADGRSDHTLTSPGGGGKAK 2404
Db        2093 RQKHEIESLYTKLGKVPP-----------AVIIPPAAPLS---GRRRRPTKS-----KGS 2133

Qy       2405 VSGRPSSRKAKSP-APGLASGDRPPSV----SSVHSEGD 2438
Db        2134 KSSRSSSLGNKSPQLSGNLSGQSAASVLHPQQTLHPPGN 2172
```

RESULT 54
US-10-336-472-230
; Sequence 230, Application US/10336472
; Publication No. US20040043929A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David W.
; APPLICANT: Ballinger, Robert A.
; APPLICANT: Baumgartner, Jason C.
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Chant, John S.
; APPLICANT: Berghs, Constance
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Gerlach, Valerie
; APPLICANT: Gilbert, Jennifer A.
; APPLICANT: Gunther, Erik
; APPLICANT: Gorman, Linda
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Ji, Weizhen
; APPLICANT: Li, Li

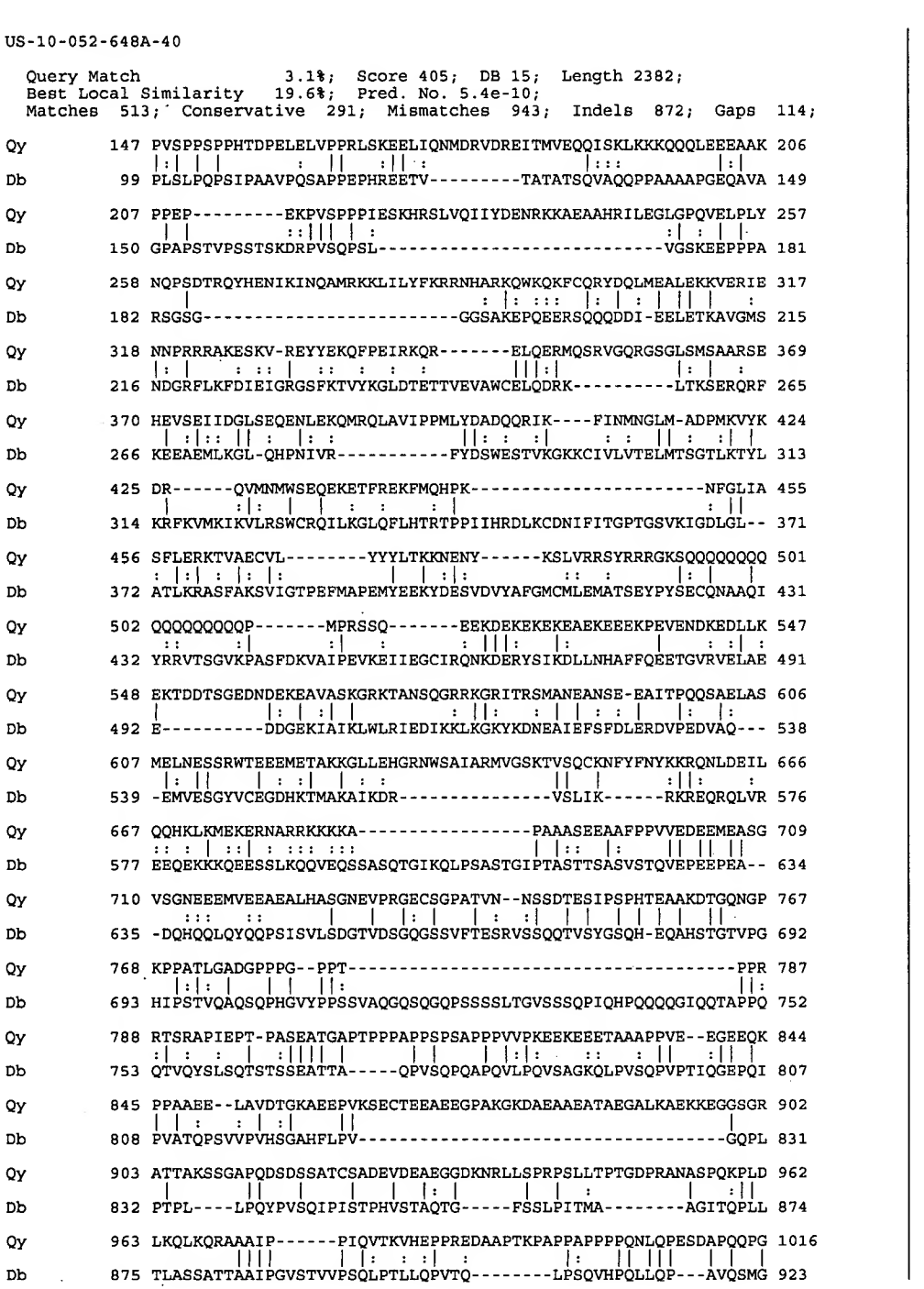
```
;  APPLICANT:  Liu, Xiaohong
;  APPLICANT:  Miller, Charles E.
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  Padigaru, Muralidhara
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Rastelli, Luca
;  APPLICANT:  MacDougall, John R.
;  APPLICANT:  Mishra, Vishnu
;  APPLICANT:  Pena, Carol E.A.
;  APPLICANT:  Spaderna, Steven K.
;  APPLICANT:  Shimkets, Richard A.
;  APPLICANT:  Smithson, Glennda
;  APPLICANT:  Spytek, Kimberly A.
;  APPLICANT:  Stone, David J.
;  APPLICANT:  Shenoy, Suresh G.
;  APPLICANT:  Ort, Tatiana
;  APPLICANT:  Taupier Jr, Raymond J.
;  APPLICANT:  Tchernev, Velizar T.
;  APPLICANT:  Vernet, Corine A.M.
;  APPLICANT:  Wolenc, Adam R.
;  APPLICANT:  Zerhusen, Bryan D.
;  APPLICANT:  Zhong, Mei
;  TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 21402-533C
;  CURRENT APPLICATION NUMBER: US/10/336,472
;  CURRENT FILING DATE:  2003-01-03
;  PRIOR APPLICATION NUMBER: 09/746,491
;  PRIOR FILING DATE: 2000-12-20
;  PRIOR APPLICATION NUMBER: 10/005,041
;  PRIOR FILING DATE: 2001-12-04
;  PRIOR APPLICATION NUMBER: 10/023,681
;  PRIOR FILING DATE: 2001-12-18
;  PRIOR APPLICATION NUMBER: 10/024,212
;  PRIOR FILING DATE: 2001-12-18
;  PRIOR APPLICATION NUMBER: 10/055,569
;  PRIOR FILING DATE: 2001-10-26
;  PRIOR APPLICATION NUMBER: 10/080,334
;  PRIOR FILING DATE: 2002-02-21
;  PRIOR APPLICATION NUMBER: 10/092,900
;  PRIOR FILING DATE: 2002-03-07
;  PRIOR APPLICATION NUMBER: 10/136,826
;  PRIOR FILING DATE: 2002-05-01
;  PRIOR APPLICATION NUMBER: 10/236,417
;  PRIOR FILING DATE: 2002-09-06
;  PRIOR APPLICATION NUMBER: 60/345,092
;  PRIOR FILING DATE: 2002-01-04
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 230
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 230
;   LENGTH: 2382
;   TYPE: PRT
;   ORGANISM: WNK1
US-10-336-472-230

  Query Match             3.1%;  Score 405;  DB 12;  Length 2382;
  Best Local Similarity   19.6%;  Pred. No. 5.4e-10;
  Matches  513;  Conservative  291;  Mismatches  943;  Indels  872;  Gaps  114;

Qy        147 PVSPPSPPHTDPELELVPPRLSKEELIQNMDRVDREITMVEQQISKLKKKQQQLEEEAAK 206
              |:| | |      :  ||   :|| :             |:::          |:|
Db         99 PLSLPQPSIPAAVPQSAPPEPHREETV---------TATATSQVAQQPPAAAAPGEQAVA 149

Qy        207 PPEP---------EKPVSPPPIESKHRSLVQIIYDENRKKAEAAHRILEGLGPQVELPLY 257
              | |          ::||| | :                    :| : | |
Db        150 GPAPSTVPSSTSKDRPVSQPSL-------------------------VGSKEEPPPA 181

Qy        258 NQPSDTRQYHENIKINQAMRKKLILYFKRRNHARKQWKQKFCQRYDQLMEALEKKVERIE 317
                  |                          : |: :::  |: | : | ||| |   :
Db        182 RSGSG-----------------------GGSAKEPQEERSQQQDDI-EELETKAVGMS 215

Qy        318 NNPRRRAKESKV-REYYEKQFPEIRKQR-------ELQERMQSRVGQRGSGLSMSAARSE 369
              |: |    : :: |  ::  :  :  :        |||:|          |: |  :

Db        216 NDGRFLKFDIEIGRGSFKTVYKGLDTETTVEVAWCELQDRK----------LTKSERQRF 265

Qy        370 HEVSEIIDGLSEQENLEKQMRQLAVIPPMLYDADQQRIK----FINMNGLM-ADPMKVYK 424
              | :|:: || :  |: :                ||: :  :|     : :  || : :| | |
Db        266 KEEAEMLKGL-QHPNIVR----------FYDSWESTVKGKKCIVLVTELMTSGTLKTYL 313

Qy        425 DR------QVMNMWSEQEKETFREKFMQHPK----------------------NFGLIA 455
              |      :|:  |  |  :  :    : |                        : ||
Db        314 KRFKVMKIKVLRSWCRQILKGLQFLHTRTPPIIHRDLKCDNIFITGPTGSVKIGDLGL-- 371

Qy        456 SFLERKTVAECVL--------YYYLTKKNENY------KSLVRRSYRRRGKSQQQQQQQQ 501
              : |:| : |: |:          |  | :|:          ::  :       |: |  |
Db        372 ATLKRASFAKSVIGTPEFMAPEMYEEKYDESVDVYAFGMCMLEMATSEYPYSECQNAAQI 431

Qy        502 QQQQQQQQQP-------MPRSSQ-------EEKDEKEKEKEAEKEEEKPEVENDKEDLLK 547
              ::     :|       :|  :       : |||:   |:          |     : :| :
Db        432 YRRVTSGVKPASFDKVAIPEVKEIIEGCIRQNKDERYSIKDLLNHAFFQEETGVRVELAE 491

Qy        548 EKTDDTSGEDNDEKEAVASKGRKTANSQGRRKGRITRSMANEANSE-EAITPQQSAELAS 606
              |          |: | :| |         : ||:   : | | : : |   |:  |:
Db        492 E----------DDGEKIAIKLWLRIEDIKKLKGKYKDNEAIEFSFDLERDVPEDVAQ--- 538

Qy        607 MELNESSRWTEEEMETAKKGLLEHGRNWSAIARMVGSKTVSQCKNFYFNYKKRQNLDEIL 666
              |: ||    | : :|  | : :                    ||  |      :||:    :
Db        539 -EMVESGYVCEGDHKTMAKAIKDR--------------VSLIK------RKREQRQLVR 576

Qy        667 QQHKLKMEKERNARRKKKKA-----------------PAAASEEAAFPPVVEDEEMEASG 709
              :: : | ::| : ::: :::                 | |::  |:    || || ||
Db        577 EEQEKKKQEESSLKQQVEQSSASQTGIKQLPSASTGIPTASTTSASVSTQVEPEEPEA-- 634

Qy        710 VSGNEEEMVEEAEALHASGNEVPRGECSGPATVN--NSSDTESIPSPHTEAAKDTGQNGP 767
                 :::   ::       |   |  |:|   | :  :|  | |  | | | | |  ||
Db        635 -DQHQQLQYQQPSISVLSDGTVDSGQGSSVFTESRVSSQQTVSYGSQH-EQAHSTGTVPG 692

Qy        768 KPPATLGADGPPPG--PPT-------------------------------------PPR 787
              |:|: |   | |  ||:                                       ||:
Db        693 HIPSTVQAQSQPHGVYPPSSVAQGQSQGQPSSSSLTGVSSSQPIQHPQQQQGIQQTAPPQ 752

Qy        788 RTSRAPIEPT-PASEATGAPTPPPAPPSPSAPPPVVPKEEKEEETAAAPPVE--EGEEQK 844
              :| :  :  |  :|||| |      | |    | |:|:    ::   : ||  :|| |
Db        753 QTVQYSLSQTSTSSEATTA-----QPVSQPQAPQVLPQVSAGKQLPVSQPVPTIQGEPQI 807

Qy        845 PPAAEE--LAVDTGKAEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGSGR 902
              | | :   : | :|    ||                                       |
Db        808 PVATQPSVVPVHSGAHFLPV-----------------------------------GQPL 831

Qy        903 ATTAKSSGAPQDSDSSATCSADEVDEAEGGDKNRLLSPRPSLLTPTGDPRANASPQKPLD 962
              |       ||   |    |  |  |: |      |  |  :         |   :||
Db        832 PTPL----LPQYPVSQIPISTPHVSTAQTG-----FSSLPITMA--------AGITQPLL 874

Qy        963 LKQLKQRAAAIP------PIQVTKVHEPPREDAAPTKPAPPAPPPPQNLQPESDAPQQPG 1016
                    ||||      | |:  : :|  :         |:   || |||   | | |
Db        875 TLASSATTAAIPGVSTVVPSQLPTLLQPVTQ--------LPSQVHPQLLQP---AVQSMG 923

Qy       1017 SSPRGKSRSPAPPADKEAFAAEAQKLPGDPPCWTSGLPFPVPPREVIKASPHAPDPSAFS 1076
                      ||:   |||    ||    | |  :||:                  |   :
Db        924 I-----------PANL-GQAAEVPLSSGD--VLYQGFPPRLPPQY----------PGDSN 959

Qy       1077 YAPPGHPLPLGLHDTARPVL--PRPPTISNPPPLISSAKHPSV---LERQIGAISQGMSV 1131
              ||  :   : :| |   ||  | |  :   |    :   | |   |   :| :  | |
Db        960 IAPSSNVASVCIHST---VLSPPMPTEVLATPGYFPTVVQPYVESNLLVPMGGV--GGQV 1014

Qy       1132 QLHVPYSEHAKAPVGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPESLGV----- 1186
              |:  |    |:||       :          |    ||          | | | :|
Db       1015 QVSQPGGSLAQAPTTSSQQAV--------LESTQGV--------SQVAPAEPVAVAQPQA 1058

Qy       1187 --PTAQEASVLRGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHGTPADVLYKGTIT 1244
                ||   :||   :|   |  | :: |  :  ||| |                ::|  |
Db       1059 TQPTTLASSV--DSAHSDVASG-MSDG--NENVPSSSG--------------RHEGRTT 1098

Qy       1245 RIIGEDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKE------------ 1292
              :     |      | : | ||  ::     ||           | :|  |
Db       1099 KRHYRKSVRSRSRHEKTSRPKLRILNVSNKGD--------RVVECQLETHNRKMVTFKFD 1150
```

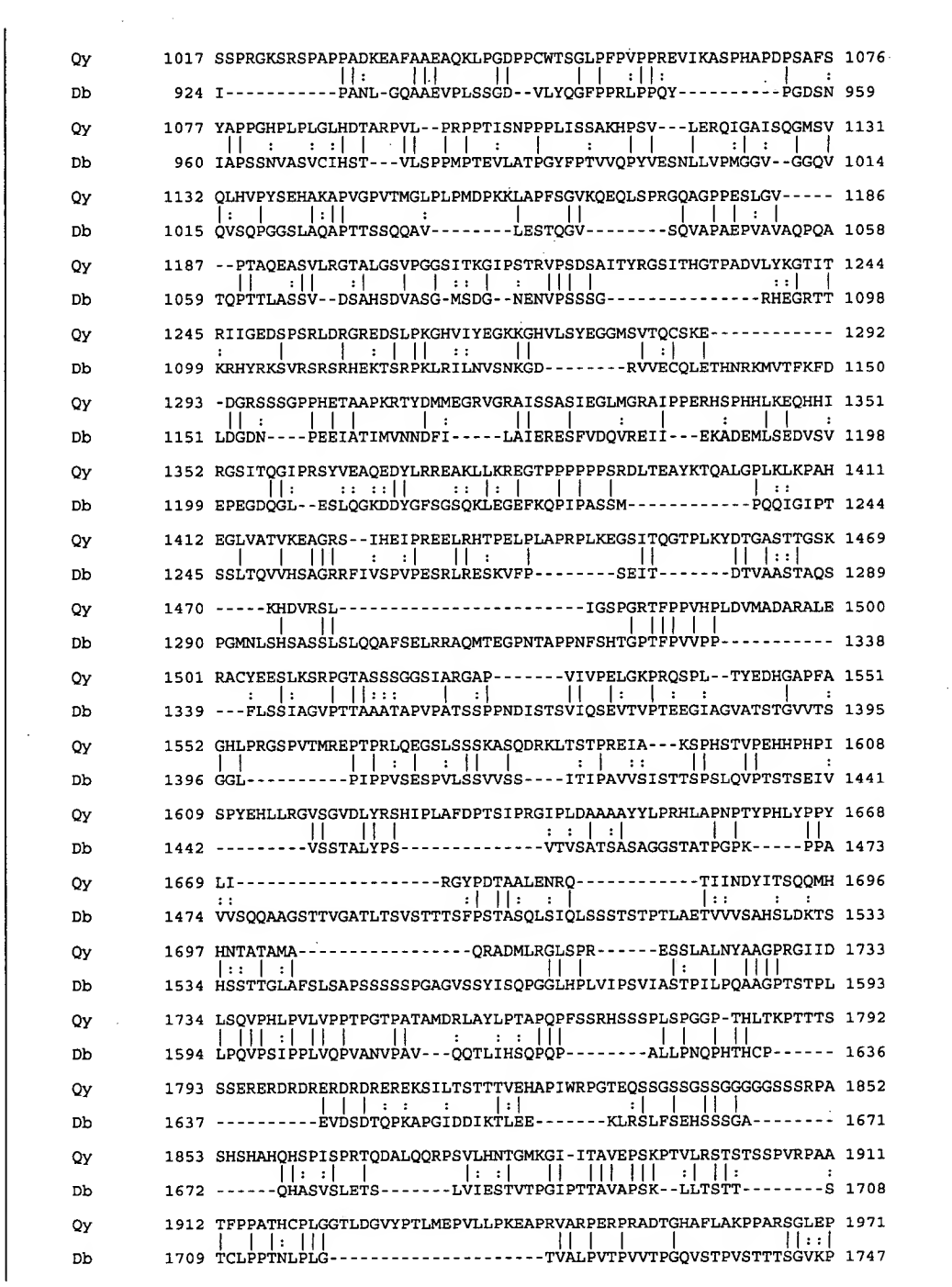
```
Qy       1293 -DGRSSSGPPHETAAPKRTYDMMEGRVGRAISSASIEGLMGRAIPPERHSPHHLKEQHHI 1351
Db       1151 LDGDN----PEEIATIMVNNDFI-----LAIERESFVDQVREII---EKADEMLSEDVSV 1198

Qy       1352 RGSITQGIPRSYVEAQEDYLRREAKLLKREGTPPPPPPSRDLTEAYKTQALGPLKLKPAH 1411
Db       1199 EPEGDQGL--ESLQGKDDYGFSGSQKLEGEFKQPIPASSM-----------PQQIGIPT 1244

Qy       1412 EGLVATVKEAGRS--IHEIPREELRHTPELPLAPRPLKEGSITQGTPLKYDTGASTTGSK 1469
Db       1245 SSLTQVVHSAGRRFIVSPVPESRLRESKVFP--------SEIT-------DTVAASTAQS 1289

Qy       1470 -----KHDVRSL------------------------IGSPGRTFPPVHPLDVMADARALE 1500
Db       1290 PGMNLSHSASSLSLQQAFSELRRAQMTEGPNTAPPNFSHTGPTFPVVPP----------- 1338

Qy       1501 RACYEESLKSRPGTASSSGGSIARGAP-------VIVPELGKPRQSPL--TYEDHGAPFA 1551
Db       1339 ---FLSSIAGVPTTAAATAPVPATSSPPNDISTSVIQSEVTVPTEEGIAGVATSTGVVTS 1395

Qy       1552 GHLPRGSPVTMREPTPRLQEGSLSSSKASQDRKLTSTPREIA---KSPHSTVPEHHPHPI 1608
Db       1396 GGL----------PIPPVSESPVLSSVVSS----ITIPAVVSISTTSPSLQVPTSTSEIV 1441

Qy       1609 SPYEHLLRGVSGVDLYRSHIPLAFDPTSIPRGIPLDAAAAYYLPRHLAPNPTYPHLYPPY 1668
Db       1442 ---------VSSTALYPS-------------VTVSATSASAGGSTATPGPK-----PPA 1473

Qy       1669 LI--------------------RGYPDTAALENRQ------------TIINDYITSQQMH 1696
Db       1474 VVSQQAAGSTTVGATLTSVSTTTSFPSTASQLSIQLSSSTSTPTLAETVVVSAHSLDKTS 1533

Qy       1697 HNTATAMA-----------------QRADMLRGLSPR------ESSLALNYAAGPRGIID 1733
Db       1534 HSSTTGLAFSLSAPSSSSSPGAGVSSYISQPGGLHPLVIPSVIASTPILPQAAGPTSTPL 1593

Qy       1734 LSQVPHLPVLVPPTPGTPATAMDRLAYLPTAPQPFSSRHSSSPLSPGGP-THLTKPTTTS 1792
Db       1594 LPQVPSIPPLVQPVANVPAV---QQTLIHSQPQP--------ALLPNQPHTHCP------ 1636

Qy       1793 SSERERDRDRERDRDREREKSILTSTTTVEHAPIWRPGTEQSSGSSGSSGGGGGSSSRPA 1852
Db       1637 ----------EVDSDTQPKAPGIDDIKTLEE-------KLRSLFSEHSSSGA-------- 1671

Qy       1853 SHSHAHQHSPISPRTQDALQQRPSVLHNTGMKGI-ITAVEPSKPTVLRSTSTSSPVRPAA 1911
Db       1672 ------QHASVSLETS-------LVIESTVTPGIPTTAVAPSK--LLTSTT--------S 1708

Qy       1912 TFPPATHCPLGGTLDGVYPTLMEPVLLPKEAPRVARPERPRADTGHAFLAKPPARSGLEP 1971
Db       1709 TCLPPTNLPLG--------------------TVALPVTPVVTPGQVSTPVSTTTSGVKP 1747

Qy       1972 ASSPSKGSEPRPLVPPVSGHATIARTPAKNLAPHHASPDPPAPPASASDPHREKTQS-KP 2030
Db       1748 GTAPSKPPLTKAPVLPVGTELPAGTLPSEQL--------PPFPGPSL-------TQSQQP 1792

Qy       2031 FSIQELELRSLGYHGSSYSPEGVEPVSPVSSPSLTHDKGLPKHLEELDKSHLEGELRPKQ 2090
Db       1793 LEDLDAQLR------RTLSPEIITVTSAV------------------------------- 1815

Qy       2091 PGPVKLGGEAAHLPHLRPLPE---SQPSSSPLLQTAPGV----KGHQRVVTLAQHISEVI 2143
Db       1816 -GPVSMAAPTA-ITEAGTQPQKGVSQVKEGPVLATSSGAGVFKMGRFQVSVAA------- 1866

Qy       2144 TQDYTRHHPQQLSAPLPAPLYSFPGASCPVLDLRRPPSDLYLPPP----------DHGAP 2193
Db       1867 --DGAQKEGKNKSEDAKSVHFESSTSESSVLSSSSPESTLVKPEPNGITIPGISSDVPES 1924

Qy       2194 ARGSPHSEGGKRSPEPNKTS--------------VLGGGEDGIE------PVSPPEGMTE 2233
Db       1925 AHKTTASEAKSDTGQPTKVGRFQVTTTANKVGRFSVSKTEDKITDTKKEGPVASPPFMDL 1984

Qy       2234 PGHSRSAVYPLLYRDGEQTEPSRMGSKSPGNTSQP-PAFFSKLTESNSAMVKSKKQEINK 2292
Db       1985 EQAVLPAVIPKKEKP-ELSEPSHLN----GPSSDPEAAFLSRDVDDGSGSPHSPHQLSSK 2039

Qy       2293 KLNTHNRNEPEYNISQPGTEIFNMPAITGTGLMTYRSQAVQ--------EHASTNMGLEA 2344
Db       2040 SL-------PSQNLSQSLSNSFNSSYMSSDNESDIEDEDLKLELRRLRDKHLKEIQDLQS 2092

Qy       2345 IIRKALMGKYDQWEESPPLSANAFNPLNASASLPAAMPITAADGRSDHTLTSPGGGGKAK 2404
Db       2093 RQKHEIESLYTKLGKVPP-----------AVIIPPAAPLS---GRRRRPTKS-----KGS 2133

Qy       2405 VSGRPSSRKAKSP-APGLASGDRPPSV----SSVHSEGD 2438
Db       2134 KSSRSSSLGNKSPQLSGNLSGQSAASVLHPQQTLHPPGN 2172
```

RESULT 55
US-10-052-648A-40
; Sequence 40, Application US/10052648A
; Publication No. US20040005558A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David
; APPLICANT: Burgess, Catherine
; APPLICANT: Casman, Stacie
; APPLICANT: Colman, Steven
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Gerlach, Valerie
; APPLICANT: Gunther, Erik
; APPLICANT: Kekuda, Ramesh
; APPLICANT: MacDougall, John R.
; APPLICANT: Mehraban, Fuad
; APPLICANT: Patturajan, Meera
; APPLICANT: Rothenberg, Mark
; APPLICANT: Shimkets, Richard
; APPLICANT: Smithson, Glennda
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Stone, David J.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Zerhusen, Bryan D.
; TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
; TITLE OF INVENTION: USING THE SAME
; FILE REFERENCE: 21402-250 (CURA-550)
; CURRENT APPLICATION NUMBER: US/10/052,648A
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: 60/262,454
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/272,920
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/284,549
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/303,229
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/262,892
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/263,605
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/269,098
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: 60/264,159
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/271,855
; PRIOR FILING DATE: 2001-02-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 2382
; TYPE: PRT
; ORGANISM: Homo sapiens

US-10-052-648A-40

Query Match 3.1%; Score 405; DB 15; Length 2382;
Best Local Similarity 19.6%; Pred. No. 5.4e-10;
Matches 513; Conservative 291; Mismatches 943; Indels 872; Gaps 114;





```
Qy       1972 ASSPSKGSEPRPLVPPVSGHATIARTPAKNLAPHHASPDPPAPPASASDPHREKTQS-KP 2030
Db       1748 GTAPSKPPLTKAPVLPVGTELPAGTLPSEQL--------PPFPGPSL-------TQSQQP 1792

Qy       2031 FSIQELELRSLGYHGSSYSPEGVEPVSPVSSPSLTHDKGLPKHLEELDKSHLEGELRPKQ 2090
Db       1793 LEDLDAQLR------RTLSPEIITVTSAV------------------------------ 1815

Qy       2091 PGPVKLGGEAAHLPHLRPLPE---SQPSSSPLLQTAPGV----KGHQRVVTLAQHISEVI 2143
Db       1816 -GPVSMAAPTA-ITEAGTQPQKGVSQVKEGPVLATSSGAGVFKMGRFQVSVAA------- 1866

Qy       2144 TQDYTRHHPQQLSAPLPAPLYSFPGASCPVLDLRRPPSDLYLPPP----------DHGAP 2193
Db       1867 --DGAQKEGKNKSEDAKSVHFESSTSESSVLSSSSPESTLVKPEPNGITIPGISSDVPES 1924

Qy       2194 ARGSPHSEGGKRSPEPNKTS--------------VLGGGEDGIE------PVSPPEGMTE 2233
Db       1925 AHKTTASEAKSDTGQPTKVGRFQVTTTANKVGRFSVSKTEDKITDTKKEGPVASPPFMDL 1984

Qy       2234 PGHSRSAVYPLLYRDGEQTEPSRMGSKSPGNTSQP-PAFFSKLTESNSAMVKSKKQEINK 2292
Db       1985 EQAVLPAVIPKKEKP-ELSEPSHLN----GPSSDPEAAFLSRDVDDGSGSPHSPHQLSSK 2039

Qy       2293 KLNTHNRNEPEYNISQPGTEIFNMPAITGTGLMTYRSQAVQ--------EHASTNMGLEA 2344
Db       2040 SL-------PSQNLSQSLSNSFNSSYMSSDNESDIEDEDLKLELRRLRDKHLKEIQDLQS 2092

Qy       2345 IIRKALMGKYDQWEESPPLSANAFNPLNASASLPAAMPITAADGRSDHTLTSPGGGGKAK 2404
Db       2093 RQKHEIESLYTKLGKVPP-----------AVIIPPAAPLS---GRRRRPTKS-----KGS 2133

Qy       2405 VSGRPSSRKAKSP-APGLASGDRPPSV----SSVHSEGD 2438
Db       2134 KSSRSSSLGNKSPQLSGNLSGQSAASVLHPQQTLHPPGN 2172
```

```
RESULT 56
US-10-124-557-14
; Sequence 14, Application US/10124557
; Publication No. US20020137894A1
;   GENERAL INFORMATION:
;        APPLICANT: Turner, Katherine
;                   Clark, Stephen C.
;                   Jacobs, Kenneth
;                   Hewick, Rodney M.
;                   Gesner, Thomas G.
;        TITLE OF INVENTION: Megakaryocyte Stimulating Factors
;        NUMBER OF SEQUENCES: 143
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Genetics Institute, Inc.
;             STREET: 87 CambridgePark Drive
;             CITY: Cambridge
;             STATE: Massachusetts
;             COUNTRY: U.S.A.
;             ZIP: 02140
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/124,557
;             FILING DATE: 16-Apr-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US 07/643,502
;             FILING DATE: 18-JAN-1991
;             APPLICATION NUMBER: US 07/546,114
;             FILING DATE: 29-JUN-1990
;             APPLICATION NUMBER: US 07/457,196
;             FILING DATE: 29-DEC-1989
;             APPLICATION NUMBER: US 07/390,901
;             FILING DATE: 08-AUG-1989
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Cserr, Luann
;             REGISTRATION NUMBER: 31,822
;             REFERENCE/DOCKET NUMBER: GI 5190
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (617)876-1170
;             TELEFAX: (617)876-5851
;   INFORMATION FOR SEQ ID NO: 14:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 941 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-10-124-557-14

  Query Match             3.0%;  Score 400.5;  DB 13;  Length 941;
  Best Local Similarity   21.4%;  Pred. No. 2.9e-10;
  Matches  259;  Conservative  116;  Mismatches  484;  Indels  351;  Gaps   55;

Qy        520 KDEKEKEKEAEKEEEKPEVENDKEDLLKEKTDDTSGEDNDEKEAVASKGRKTANSQGRRK 579
Db          2 KDNK-KNRTKKKPTPKPPVVDEA----------GSGLDNGDFKVTTPDTSTTQHNKVSTS 50

Qy        580 GRITRSMANEANSEEAITP-QQSAELASMELNESSRWTEEEMETAKKGLLEHGRNWSAIA 638
Db         51 PKIT--TAKPINPRPSLPPNSDTSKETSLTVNKETTVETKETTTTNKQTSTDGKEKTTSA 108

Qy        639 RMVGSKTVSQCKNFYFNYKKRQNLDEILQQHKLKME---KERNARRKKKKAPAAASEEAA 695
Db        109 KETQSIEKTSAKDL-------APTSKVLAKPTPKAETTTKGPALTTPKEPTPTTPKEPAS 161

Qy        696 PPPVVEDEEMEASGVSGNEEEMVEEAEALHASGNEVPRGECSGPATVNNSSDTESIPSPH 755
Db        162 TTP-------------------------------KEPT-----PTTIKSAPTTPKEPAPT 185

Qy        756 TEAAKDTGQNGPKPPATLGADGPPPGPPTPPRRTSRAPIEPTPASEATGAPTPP--PAPP 813
Db        186 TTKSAPTTPKEPAPTTT-----KEPAPTTPKEPAPTTTKEPAPTT-TKSAPTTPKEPAPT 239

Qy        814 SPSAPPPVVPKEEKEEETAAAPPVEEGEEQKPPAAEELAVDTGKAEEPVKSECTEEAEEG 873
Db        240 TPKKPAPTTPKE--PAPTTPKEPTPTTPKEPAPTTKEPAPTTPKEPAPT-------APKK 290

Qy        874 PAKGKDAEAAEATAEGALKAEKKEGGSGRATTAKSSGAPQDSDSSATCSADEVDEAEGGD 933
Db        291 PAPTTPKEPAPTTPKEPAPTTTKE----PSPTTPKEPAPTTTKSAPTTTKEP-------- 338

Qy        934 KNRLLSPRPSLLTPTGDPRANASPQKPLDLKQLKQRAAAIPPIQVTKVHEP----PREDA 989
Db        339 ------------APTTTKSAPTTPKEP---------------SPTTTKEPAPTTPKEPA 370

Qy        990 --APTKPAPPAPPPPQNLQPESDAPQQPGSSPRGKSRSPAPPADKEAFAAEAQKLPGDPP 1047
Db        371 PTTPKKPAPTTPKEPAPTTPKEPAPTTTKKPAPTAPKEPAPTTPKETAPTTPKKLTPTTP 430

Qy       1048 ----CWTSGLPFPVPPREVIKASPHAPDPSAFSYAPPGHPLPLGLHDTARPVLPRPPTIS 1103
Db        431 EKLAPTTPEKPAPTTPEELAPTTPEEPTPTT-----PEEPAPT-TPKAAAPNTPKEPAPT 484

Qy       1104 NP-PPLISSAKHPSVLERQIGAISQGMSVQLHVPYSEHAKAP-----VGPVTMGLPLPMD 1157
Db        485 TPKEPAPTTPKEPA------------------PTTPKETAPTTPKGTAPTTLKEPAPTT 525

Qy       1158 PKKLAPFSGVKQEQLSPRGQAGPPESLG---VPTAQEASVLRGTALGSVPGGSITKGIPS 1214
Db        526 PKKPAP------KELAPTTTKEPTSTTSDKPAPTTP-----KGTA------------PT 561

Qy       1215 TRVPSDSAITYRGSITHGTPADVLYKGTITRIIGEDSPSRLDRGREDSLPKGHVIYEGKK 1274
Db        562 T--PKEPAPT-----TPKEPAPTTPKGTAPTTLKEPAPTTPKKPAPKELAP--------- 605
```

```
Qy       1275 GHVLSYEGGMSVTQCSKEDGRSSSGPPHETA--APKRTYDMMEGRVGRAISSASIEGLMG 1332
                  : :|  | |       : :   | |||   ||        :
Db        606 ---TTTKGPTSTT-----SDKPAPTTPKETAPTTPKEPAPTTPKKPAPTTPETP------ 651

Qy       1333 RAIPP---ERHSPHHLKEQHHIRGSITQGIPRSYVEAQEDYLRREAKLLKREGTP----- 1384
                ||   | :|   ||   |  | : |             |  | ||
Db        652 ---PPTTSEVSTPTTTKEPTTIHKSPDESTPE---------------LSAEPTPKALEN 692

Qy       1385 PPPPPSRDLTEAYKTQALGPLKLKPAHEGLVATVKEAGRSIHEIPREELRHTPE-LPLAP 1443
                |  |    |:        |  ||    :  | |:      :    :|| |||    ||
Db        693 SPKEPGVPTTKT-------PAATKPE---MTTTAKD------KTTERDLRTTPETTTAAP 736

Qy       1444 RPLKEGSIT----------------------QGTPLKYDTGAS-------TTGSKKHDVR 1474
                :  || : |                    || |  |  :       ||  |
Db        737 KMTKETATTTEKTTESKITATTTQVTSTTTQDTTPFKITTLKTTTLAPKVTTTKKTITTT 796

Qy       1475 SLIGSPGRTFPPVHPLDVMADARALERACYEESLKSRPGTASSSGGS-IARGAPVIVPEL 1533
                ::  |  |         . |:  :||  ::   :|  : :    :   |  :| :
Db        797 EIMNKPEET-----------AKPKDRATNSKATTPKPQKPTKAPKKPTSTKKPKTMPRV 844

Qy       1534 GKPRQSPLTYEDHGAPFAGHLPRGSPVTMRE--PTPRLQEGSLSSSKASQDRKLTSTPRE 1591
                 ||: :|              ||    || |  || |: |  | :          |: | :
Db        845 RKPKTTP-------------TPRKMTSTMPELNPTSRIAEAMLQT---------TTRPNQ 882

Qy       1592 IAKSPHSTVPEHHPHPISPYEHLLRGVSGVD---LYRSHIPLAFDPTSIPRGIPLDAAAA 1648
                   :|:| : |     ::|       | |      | | |:   | |  |    :|
Db        883 ---TPNSKLVE-----VNPKSEDAGGAEGETPHMLLRPHV---FMPEVTP---DMD---- 924

Qy       1649 YYLPRHLAPN 1658
                 ||||   ||
Db        925 -YLPR--VPN 931


RESULT 57
US-10-124-557-84
; Sequence 84, Application US/10124557
; Publication No. US20020137894A1
;   GENERAL INFORMATION:
;        APPLICANT: Turner, Katherine
;                   Clark, Stephen C.
;                   Jacobs, Kenneth
;                   Hewick, Rodney M.
;                   Gesner, Thomas G.
;        TITLE OF INVENTION: Megakaryocyte Stimulating Factors
;        NUMBER OF SEQUENCES: 143
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Genetics Institute, Inc.
;             STREET: 87 CambridgePark Drive
;             CITY: Cambridge
;             STATE: Massachusetts
;             COUNTRY: U.S.A.
;             ZIP: 02140
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/124,557
;             FILING DATE: 16-Apr-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US 07/643,502
;             FILING DATE: 18-JAN-1991
;             APPLICATION NUMBER: US 07/546,114
;             FILING DATE: 29-JUN-1990
;             APPLICATION NUMBER: US 07/457,196
;             FILING DATE: 29-DEC-1989
;             APPLICATION NUMBER: US 07/390,901
;             FILING DATE: 08-AUG-1989
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Cserr, Luann
;             REGISTRATION NUMBER: 31,822
;             REFERENCE/DOCKET NUMBER: GI 5190
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (617)876-1170
;             TELEFAX: (617)876-5851
;   INFORMATION FOR SEQ ID NO: 84:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 1022 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 84:
US-10-124-557-84

  Query Match             3.0%;  Score 400.5;  DB 13;  Length 1022;
  Best Local Similarity   21.4%;  Pred. No. 3.2e-10;
  Matches  259;  Conservative  116;  Mismatches  484;  Indels  351;  Gaps   55;

Qy        520 KDEKEKEKEAEKEEEKPEVENDKEDLLKEKTDDTSGEDNDEKEAVASKGRKTANSQGRRK 579
                 || | | :  :|   || | ::            || || : :         | :::
Db         83 KDNK-KNRTKKKPTPKPPVVDEA----------GSGLDNGDFKVTTPDTSTTQHNKVSTS 131

Qy        580 GRITRSMANEANSEEAITP-QQSAELASMELNESSRWTEEEMETAKKGLLEHGRNWSAIA 638
                :||   |   |   :: |   :::  |: :|: :    :|  |  |      |:  :  |
Db        132 PKIT--TAKPINPRPSLPPNSDTSKETSLTVNKETTVETKETTTTNKQTSTDGKEKTTSA 189

Qy        639 RMVGSKTVSQCKNFYFNYKKRQNLDEILQQHKLKME---KERNARRKKKKAPAAASEEAA 695
                :   |  :  |:                ::| :   | |   |       |:  |   | |:
Db        190 KETQSIEKTSAKDL-------APTSKVLAKPTPKAETTTKGPALTTPKEPTPTTPKEPAS 242

Qy        696 FPPVVEDEEMEASGVSGNEEEMVEEAEALHASGNEVPRGECSGPATVNNSSDTESIPSPH 755
                  |                                : |      | |: ::  |   |:|
Db        243 TTP-------------------------------KEPT-----PTTIKSAPTTPKEPAPT 266

Qy        756 TEAAKDTGQNGPKPPATLGADGPPPGPPTPPRRTSRAPIEPTPASEATGAPTPP--PAPP 813
                |  :  |    | |  |        | | ||          || | :    ||| |  |||
Db        267 TTKSAPTTPKEPAPTTT-----KEPAPTTPKEPAPTTTKEPAPTT-TKSAPTTPKEPAPT 320

Qy        814 SPSAPPPVVPKEEKEEETAAAPPVEEGEEQKPPAAEELAVDTGKAEEPVKSECTEEAEEG 873
                :|  | |  |||      |    |      ::  |  :| |  | |   |        | :
Db        321 TPKKPAPTTPKE--PAPTTPKEPTPTTPKEPAPTTKEPAPTTPKEPAPT-------APKK 371

Qy        874 PAKGKDAEAAEATAEGALKAEKKEGGSGRATTAKSSGAPQDSDSSATCSADEVDEAEGGD 933
                ||     | |  | :       ||     : |     ||  : |: | : :
Db        372 PAPTTPKEPAPTTPKEPAPTTTKE----PSPTTPKEPAPTTTKSAPTTTKEP-------- 419

Qy        934 KNRLLSPRPSLLTPTGDPRANASPQKPLDLKQLKQRAAAIPPIQVTKVHEP----PREDA 989
                          ||     | :|::|                  |   ||     |:| |
Db        420 -----------APTTTKSAPTTPKEP---------------SPTTTKEPAPTTPKEPA 451

Qy        990 --APTKPAPPAPPPPQNLQPESDAPQQPGSSPRGKSRSPAPPADKEAFAAEAQKLPGDPP 1047
                  | ||||  |  |    |:  ||            : |||   ||      :||    |
Db        452 PTTPKKPAPTTPKEPAPTTPKEPAPTTTKKPAPTAPKEPAPTTPKETAPTTPKKLTPTTP 511

Qy       1048 ----CWTSGLPFPVPPREVIKASPHAPDPSAFSYAPPGHPLPLGLHDTARPVLPRPPTIS 1103
                    |   | |  | |:   :|  | |:       |  | |       | |  |: |  :
Db        512 EKLAPTTPEKPAPTTPEELAPTTPEEPTPTT-----PEEPAPT-TPKAAAPNTPKEPAPT 565

Qy       1104 NP-PPLISSAKHPSVLERQIGAISQGMSVQLHVPYSEHAKAP-----VGPVTMGLPLPMD 1157
                |  |  :: | |:                    | :    ||        | ||:  | |
Db        566 TPKEPAPTTPKEPA-----------------PTTPKETAPTTPKGTAPTTLKEPAPTT 606

Qy       1158 PKKLAPFSGVKQEQLSPRGQAGPPESLG---VPTAQEASVLRGTALGSVPGGSITKGIPS 1214
                |||  ||      ::|:|      |  :      ||        :|||            |:
Db        607 PKKPAP------KELAPTTTKEPTSTTSDKPAPTTP-----KGTA-------------PT 642

Qy       1215 TRVPSDSAITYRGSITHGTPADVLYKGTITRIIGEDSPSRLDRGREDSLPKGHVIYEGKK 1274
                |  | : |.|     |   ||    |||    : | :|:   :      |
Db        643 T--PKEPAPT-----TPKEPAPTTPKGTAPTTLKEPAPTTPKKPAPKELAP--------- 686

Qy       1275 GHVLSYEGGMSVTQCSKEDGRSSSGPPHETA--APKRTYDMMEGRVGRAISSASIEGLMG 1332
                  : :|  | |       : :   | |||   ||        :
Db        687 ---TTTKGPTSTT-----SDKPAPTTPKETAPTTPKEPAPTTPKKPAPTTPETP------ 732
```

```
Qy       1333 RAIPP---ERHSPHHLKEQHHIRGSITQGIPRSYVEAQEDYLRREAKLLKREGTP----- 1384
Db        733 ---PPTTSEVSTPTTTKEPTTIHKSPDESTPE---------------LSAEPTPKALEN 773

Qy       1385 PPPPPSRDLTEAYKTQALGPLKLKPAHEGLVATVKEAGRSIHEIPREELRHTPE-LPLAP 1443
Db        774 SPKEPGVPTTKT-------PAATKPE---MTTTAKD------KTTERDLRTTPETTTAAP 817

Qy       1444 RPLKEGSIT----------------------QGTPLKYDTGAS-------TTGSKKHDVR 1474
Db        818 KMTKETATTTEKTTESKITATTTQVTSTTTQDTTPFKITTLKTTTLAPKVTTTKKTITTT 877

Qy       1475 SLIGSPGRTFPPVHPLDVMADARALERACYEESLKSRPGTASSSGGS-IARGAPVIVPEL 1533
Db        878 EIMNKPEET-----------AKPKDRATNSKATTPKPQKPTKAPKKPTSTKKPKTMPRV 925

Qy       1534 GKPRQSPLTYEDHGAPFAGHLPRGSPVTMRE--PTPRLQEGSLSSSKASQDRKLTSTPRE 1591
Db        926 RKPKTTP-------------TPRKMTSTMPELNPTSRIAEAMLQT---------TTRPNQ 963

Qy       1592 IAKSPHSTVPEHHPHPISPYEHLLRGVSGVD---LYRSHIPLAFDPTSIPRGIPLDAAAA 1648
Db        964 ---TPNSKLVE-----VNPKSEDAGGAEGETPHMLLRPHV---FMPEVTP---DMD---- 1005

Qy       1649 YYLPRHLAPN 1658
Db       1006 -YLPR--VPN 1012
```

RESULT 58
US-10-216-705-21
; Sequence 21, Application US/10216705
; Publication No. US20030096973A1
; GENERAL INFORMATION:
; APPLICANT: Meristem Therapeutics, S.A.
; TITLE OF INVENTION: Recombinant Collagens and Derived Proteins Produced by Plants, Me
; TITLE OF INVENTION: obtaining Such and Their Uses
; FILE REFERENCE: 1149-3 DIV
; CURRENT APPLICATION NUMBER: US/10/216,705
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 09/331,347
; PRIOR FILING DATE: 1999-08-17
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 1464
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-216-705-21

Query Match 3.0%; Score 400.5; DB 14; Length 1464;
Best Local Similarity 21.5%; Pred. No. 4.9e-10;
Matches 380; Conservative 102; Mismatches 642; Indels 647; Gaps 83;

```
Qy        722 EALHASGNEVPRGECSGPATVNNSSDTESIPSPHTEAAK-DTGQNGPKPPA-TLGADGPP 779
Db         77 ETKNCPGAEVPEGECC-PVCPDGSESPTDQETTGVEGPKGDTGPRGPRGPAGPPGRDGIP 135

Qy        780 -----PGPPTPPRRTSRAPIEPTPASEATGAPTPP----------------------P 810
Db        136 GQPGLPGPPGPP----------------GPPGPPGLGGKLAPQLSYGYDEKSTGGISVP 178

Qy        811 APPSPSAP-----PPVVPKEEKEEETAAAPPVEEGE--------EQKPPAAEELAVDTGK 857
Db        179 GPMGPSGPRGLPGPPGAPGPQGFQ----GPPGEPGEPGASGPMGPRGPPGPPGKNGDDGE 234

Qy        858 AEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGSGRATTAKSSGAPQDSDS 917
Db        235 AGKPGR-----PGERGPPGPQGARGLPGTA--GLPGMKGHRGFSGLDGAKGDAGP----- 282

Qy        918 SATCSADEVDEAEGGDKNRLLSP----RPSLLTPTGDPRANASPQKPLDLKQLKQRAAAI 973
Db        283 ------------AGPKGEPGSPGENGAPGQMGPRGLPGERGRPGAP-------GPAGAR 322

Qy        974 PPIQVTKVHEPPREDAAPTKPA-PPAPPPPQNLQPESDAPQQPGSS--PRGKSRSPAPP- 1029
Db        323 GNDGATGAAGPP----GPTGPAGPPGFPGAVGAKGEA-GPQGPRGSEGPQGVRGEPGPPG 377

Qy       1030 ------------ADKEAFAAEAQKLPGDPPCWTSGLP-FPVPPREVIKASPHAPDPSAFS 1076
Db        378 PAGAAGPAGNPGADGQPGAKGANGAPG-----IAGAPGFPGARGPSGPQGPGGPPGPKGN 432

Qy       1077 YAPPGHPLPLGLHDTARPVLPRPPTISNPPPLISSAKHPSVLERQIGAISQGMSVQLHVP 1136
Db        433 SGEPGAPGSKG--DTGAKGEPGPVGVQGPP-------GPAGEEGKRG------------ 470

Qy       1137 YSEHAKAPVGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPESLGVPTAQEASVLR 1196
Db        471 ----ARGEPGPT--GLP-------------------GPPGERGGPGSRGFPGADGVAGPK 505

Qy       1197 GTA--LGSVPGGSITKGIP-STRVPSDSAITYRGSITHGTPADVLYKGTITRIIGEDSPS 1253
Db        506 GPAGERGS-PGPAGPKGSPGEAGRPGEAGLPGAKGLT-GSPGSPGPDGK----TGPPGPA 559

Qy       1254 RLD-----------RGREDSL----PKGHVIYEGKKGHVLSYEGGM-----SVTQCSKED 1293
Db        560 GQDGRPGPPGPPGARGQAGVMGFPGPKGAAGEPGKAG-----ERGVPGPPGAVGPAGKDG 614

Qy       1294 GRSSSGPPHETAAPKRTYDMMEGRVGRAISSASIEGLMGRAIPP-ERHSPHHLKEQHHIR 1352
Db        615 EAGAQGPP-GPAGPAGE----RGEQGPA-GSPGFQGLPGPAGPPGEAGKPGE-------- 660

Qy       1353 GSITQGIPRSYVEAQEDYLRREAKLLKREGTPPPPPPSRDLTEAYKTQALGPLKLK--PA 1410
Db        661 ----QGVPGDLGAPGPSGARGERGFPGERGVQGPPGPA-----------GPRGANGAPG 704

Qy       1411 HEGLVATVKEAGRSIHEIPREELRHTPELPLAPRPLKEGSITQGTP----LKYDTGASTT 1466
Db        705 NDGAKG---DAG---------------APGAP-----GS--QGAPGLQGMPGERGAAGL 738

Qy       1467 GSKKHD-----VRSLIGSPGR------TFP--PVHPLDVMADARALERACYEESLKSRPG 1513
Db        739 PGPKGDRGDAGPKGADGSPGKDGVRGLTGPIGPPGPAGAPGD------------KGESG 785

Qy       1514 TASSSGGSIARGAPVIVPELGKPRQSPLTYEDHGAPFAGHLPRGSPVTMREPTPRLQEGS 1573
Db        786 PSGPAGPTGARGAPGDRGEPGPP--GPAGFA--GPPGAD----GQPGAKGEPGDAGAKGD 837

Qy       1574 LSSSKASQDRKLTSTPREIAKSPHSTVPEHHPHPI----SPYEHLLRGVSGVDLYRSHIP 1629
Db        838 -------------------AGPPGPAGPAGPPGPIGNVGAPGAKGARGSAGP-------- 870

Qy       1630 LAFDPTSIPRGIPLDAAAAYYLPRHLAPNPTYPHLYPPYLIRGYPDTAALENRQTIINDY 1689
Db        871 --------PGATGFPGAAGRVGP----PGPS-GNAGPP----GPPGPAGKEGGKGPRGE- 912

Qy       1690 ITSQQMHHNTATAMAQRADML-----------RGLSPRESSLALNYAAGPRGIIDLSQVP 1738
Db        913 -----------TGPAGRPGEVGPPGPPGPAGEKGSPGADGPAGAPGTPGPQGIAGQRGVV 961

Qy       1739 HLPVLVPPT--PGTPATAMDRLAYLPTAPQPFSSRHSSSPLSPGGPTHLTKPTTTSSSER 1796
Db        962 GLPGQRGERGFPGLPGPSGE-----PGKQGPSGASGERGPPGPMGPPGLAGPPGESGRE- 1015

Qy       1797 ERDRDRERDRDRREREKSILTSTTTVEHAPIWRPGTEQSSGSSGSSGGGG-----GSSSRP 1851
Db       1016 ----------------------------GAPGAEGSPGRDGSPGAKGDRGETGPAGPP 1045

Qy       1852 ASHSHAHQHSPISPRTQDALQQRPSVLHNTGMKGIITAVEPSKPTVLRSTSTSSPVRPAA 1911
Db       1046 GAPGAPGAPGPVGP--------------AGKSGDRGETGPAGP--------AGPVGPAG 1082

Qy       1912 TFPPATHCPLGGTLDGVYPTLMEPVLLPKEAPRVARPERPRADTGH------------- 1957
Db       1083 ARGP----------------------------AGPQGPRGDKGETGEQGDRGIKGHRG 1112
```

```
Qy       1958 -AFLAKPPARSGLEPASSPSKGSEPRPLVPPVSGHATIARTPAKNLAPHHASPDPPAPPA 2016
Db       1113 FSGLQGPPGPPGSPGEQGPSGASGP--------------------------AGPRGPPG 1145

Qy       2017 SASDPHREKTQSKPFSIQELELRSLGYHGSSYSPEGVEPVSPVSSPSLTHDKGLPKHLEE 2076
Db       1146 SAGAPGKDGLNGLPGPI--------GPPGPRGRTGDAGPVGPPGPPG------------- 1184

Qy       2077 LDKSHLEGELRPKQPGPVKLGGEAAHLPHLRPLPESQPSSSPLLQTAPGVKGHQ------ 2130
Db       1185 ----------PPGPPGPPSAGFDFSFLP--------QP---------PQEKAHDGGRYYR 1217

Qy       2131 ---------------RVVTLAQHISEVITQDYTRHHPQQLSAPLPAPLYSFPGASCPVL 2174
Db       1218 ADDANVVRDRDLEVDTTLKSLSQQIENIRSPEGSRKNPAR--------------TC--R 1260

Qy       2175 DLRRPPSDLYLPPPDHGAPARGSPHSEGGKRSPEPNKTSVLGG---------GEDGIEPV 2225
Db       1261 DLKMCHSDW----------------KSGEYWIDPNQGCNLDAIKVFCNMETGETCVYPT 1303

Qy       2226 SPPEGM------TEPGHSRSAVYPLLYRDGEQTEPSRMGSKSPGNTSQPPAFFSKLTESN 2279
Db       1304 QPSVAQKNWYISKNPKDKRHVWFGESMTDGFQFEYG-------GQGSDPADVAIQLT--- 1353

Qy       2280 SAMVKSKKQEINKKLNTHNRNEPEYNISQPG 2310
Db       1354 --FLRLMSTEASQNITYHCKNSVAYMDQQTG 1382
```

RESULT 59
US-10-128-714-3305
; Sequence 3305, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wenqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; TITLE OF INVENTION: Methods of Use
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3305
; LENGTH: 1000
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-10-128-714-3305

Query Match 3.0%; Score 398; DB 14; Length 1000;
Best Local Similarity 21.8%; Pred. No. 4e-10;
Matches 215; Conservative 134; Mismatches 443; Indels 192; Gaps 36;

```
Qy        496 QQQQQQQQQQQQQQQQPMPRSSQEEKDEKEKEKEAEKEEEKPEVENDKEDLLKEKTDDTSG 555
Db        102 ESAEEDVTKPDTENSPAPDGPAEE--QSDSPPEAEQPVAEVQVPTEEQPAPKKPAPEPSS 159

Qy        556 ED----------NDEKEAVASKGRKTANSQGRRKGRITRSMANEANSEEAITPQQSAELA 605
```

```
Db        160 EEPAAPTETAENNTSEEAVAPEG-ATAD-----------IAADETTKEESTDEKAPEEPA 207

Qy        606 SMELNESSRWTEEEMETAKKGLL--EHGRNWSAIARMVGSKTVSQCKNFYFNYKKRQNLD 663
Db        208 AAESP-----VEDKQDEATKALAAEEGGTKTPAADEAVAEDAAAE------PAPEASSTD 256

Qy        664 EILQQHKLKM-----EKERNARRKKKKAPAAAS--EEAAFPPVVEDEEMEASGVSGNEEE 716
Db        257 EVAQPEATPAEAPTEEKKEDAPQEPPEQPLTEEVVEESASPESVAEEPASAEQPPPADEP 316

Qy        717 MVEEAEALHASGNEVPRGECSGPATVNNSSDTESIPSPHTEAAKDTGQNGPKPPATLGAD 776
Db        317 APEEPPAAEESAPAVPEVEEPAP---QDETPVQKAPAAEQSTAEDVTPEEPAPVEEASAK 373

Qy        777 GPPPGPPTPPRRTS---RAPIEPTPASEATGA-----PTPPPAPPSPSAPPPVVPKEEKE 828
Db        374 EPVAEEPTPGDKSSPVEEAAVEEAPPAEESAASEESTPVEEAAPTEESAAEESAPAEEAT 433

Qy        829 EETA----AAPPVEEGEEQKPPAAEELAVDTGKAEE--PVKSECTEEAE--EGPAKGKDA 880
Db        434 ESISVGEPAAADAEEGLREESAHEEPAPVEDAPAEEPAPVEEPAAEESPSVEDPAPVEEP 493

Qy        881 EAAEATAEGALKAEKKEGGSGRATTAKSSGAPQDSDSSATCSADEVDEAEGGDKNRLLSP 940
Db        494 APAEDPAAAAEEIPTEEPAPAEEVPAEEPAPVQEPAEEPTPVEKSVPEESAPAEERI--- 550

Qy        941 RPSLLTPTGD-PRANASPQKPLDLKQLKQRAAAIPPIQVTKVHEPPREDAAPTKPAPPAP 999
Db        551 ------PVEEAPAEESTPAEPAPVEEVPAEISA--PVEPAPVEEPAAEEAPPTESAREEP 602

Qy       1000 PPPQNLQPESDAPQQPGSSPRGKSRSPAP--PADKEAFAAEAQKLPGDPPCWTSGLPFPV 1057
Db        603 APME------EAPTEESASVEPASVEPAPVEPGPVEEPAAAAEEIPAEEP---AVVEEGA 653

Qy       1058 PPREVIKASPHAPDPSAFSYAPPGHPLPLGLHDTARPVLPRPPTISNPP---PLISSAKH 1114
Db        654 PVEESVPVKEAAPVDAEFSEETPD---PTPVEEISRDIDALEPLEAALPAVVPVAIAAEH 710

Qy       1115 PSVLERQIGAISQGMSVQLHVPYSEHAKAPVGPVTMGLPLPMDPKKLAPFSG-------- 1166
Db        711 ARRRKRR--SPDEGQ--------RRHSKSSSEEVRRSL-----PRKKPERSGSLLDRWNK 755

Qy       1167 ------------VKQEQLSPRGQAGPPESL--------------GVPTAQEASVLRGTA 1199
Db        756 ALEEAKRQHEEKLRQEELRSKDKFRTPEKAERDRHSSERERHKRSERSSKEEAAV----- 810

Qy       1200 LGSVPGGSITKGIPSTRVPSDSAITYRGSITHGTPADVLYKGTITRIIGEDSPSRLDRGR 1259
Db        811 -----------DVKHSRRRPERSATYEVERSSGEKAAVPIE----------PSR--RTR 846

Qy       1260 EDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPHETAAPKRTYDMMEGRVG 1319
Db        847 EHS----------------SSQGG-ERRSTSSRDGHSSVSKPRAFLKYMTAESETNGPLL 889

Qy       1320 RAISSASIEGLMGRAIPPERHSPHHLKEQHHIRGSITQGIPRSYVEAQEDYLRREAKLLK 1379
Db        890 KINGDKAAANVLGRRSSPS-HSHSH-RHSHEGRGS--DRSTSSHHAAEEEQARRERRARR 945

Qy       1380 REGTPPPPPPSRDLTEAYKTQALG 1403
Db        946 RAAEEVEQAKEREATDHHRHRHSG 969
```

RESULT 60
US-10-291-265-243
; Sequence 243, Application US/10291265
; Publication No. US20030232054A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang et al
; TITLE OF INVENTION: No. US20030232054A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-017 (785)
; CURRENT APPLICATION NUMBER: US/10/291,265

```
;  CURRENT FILING DATE:  2000-01-25
;  PRIOR APPLICATION NUMBER: 09/491,404
;  PRIOR FILING DATE: 2000-01-25
;  PRIOR APPLICATION NUMBER: 09/617,746
;  PRIOR FILING DATE: 2000-07-17
;  PRIOR APPLICATION NUMBER: 09/631,451
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: 09/633,870
;  PRIOR FILING DATE: 2000-09-15
;  NUMBER OF SEQ ID NOS: 944
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 243
;   LENGTH: 1464
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-291-265-243

  Query Match             3.0%;  Score 396.5;  DB 15;  Length 1464;
  Best Local Similarity   21.4%;  Pred. No. 7.4e-10;
  Matches  379;  Conservative  102;  Mismatches  643;  Indels  647;  Gaps   83;

Qy        722 EALHASGNEVPRGECSGPATVNNSSDTESIPSPHTEAAK-DTGQNGPKPPA-TLGADGPP 779
              |  :  | |||  |||  |  : |       :   |  | |||  ||: ||   | || |
Db         77 ETKNCPGAEVPEGECC-PVCPDGSESPTDQETTGVEGPKGDTGPRGPRGPAGPPGRDGIP 135

Qy        780 -----PGPPTPPRRTSRAPIEPTPASEATGAPTPP----------------------P 810
                  ||||  ||                  | | ||                     |
Db        136 GQPGLPGPPGPP----------------GPPGPPGLGGNFAPQLSYGYDEKSTGGISVP 178

Qy        811 APPSPSAP-----PPVVPKEEKEEETAAAPPVEEGE--------EQKPPAAEELAVDTGK 857
              |  || |      ||  |  : :     || | ||         : ||      | |:
Db        179 GPMGPSGPRGLPGPPGAPGPQGFQ----GPPGEPGEPGASGPMGPRGPPGPPGPKNGDDGE 234

Qy        858 AEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGSGRATTAKSSGAPQDSDS 917
              | :| :       | ||   : |     ||   |   |   |       ||    |
Db        235 AGKPGR-----PGERGPPGPQGARGLPGTA--GLPGMKGHRGFSGLDGAKGDAGP----- 282

Qy        918 SATCSADEVDEAEGGDKNRLLSP----RPSLLTPTGDPRANASPQKPLDLKQLKQRAAAI 973
                        | |    ||      |  : | | |      |  |          | |
Db        283 -------------AGPKGEPGSPGENGAPGQMGPRGLPGERGRPGAP-------GPAGAR 322

Qy        974 PPIQVTKVHEPPREDAAPTKPA-PPAPPPPQNLQPESDAPQQPGSS--PRGKSRSPAPP- 1029
                  |     ||     || || ||  |     : |:  || |  |  |:|    | ||
Db        323 GNDGATGAAGPP----GPTGPAGPPGFPGAVGAKGEA-GPQGPRGSEGPQGVRGEPGPPG 377

Qy       1030 ------------ADKEAFAAEAQKLPGDPPCWTSGLP-FPVPPREVIKASPHAPDPSAFS 1076
                         || :  |  |   ||      :| | ||              | |     :
Db        378 PAGAAGPAGNPGADGQPGAKGANGAPG-----IAGAPGFPGARGPSGPQGPGGPPGPKGN 432

Qy       1077 YAPPGHPLPLGLHDTARPVLPRPPTISNPPPLISSAKHPSVLERQIGAISQGMSVQLHVP 1136
                 || |   |  ||     | |  :  ||        |:  | : |
Db        433 SGEPGAPGSKG--DTGAKGEPGPVGVQGPP-------GPAGEEGKRG------------- 470

Qy       1137 YSEHAKAPVGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPESLGVPTAQEASVLR 1196
                   |:   ||   |||                  | |: | | | | | |   :  :
Db        471 ----ARGEPGPT--GLP------------------GPPGERGGPGSRGFPGADGVAGPK 505

Qy       1197 GTA--LGSVPGGSITKGIP-STRVPSDSAITYRGSITHGTPADVLYKGTITRIIGEDSPS 1253
                | |   || || :  || |     | :: :     :| |:|       |      |   |:
Db        506 GPAGERGS-PGPAGPKGSPGEAGRPGEAGLPGAKGLT-GSPGSPGPDGK----TGPPGPA 559

Qy       1254 RLD-----------RGREDSL----PKGHVIYEGKKGHVLSYEGGM-----SVTQCSKED 1293
                |            ||:   :    |||     || |     | |:     :|    |:
Db        560 GQDGRPGPPGPPGARGQAGVMGFPGPKGAAGEPGKAG-----ERGVPGPPGAVGPAGKDG 614

Qy       1294 GRSSSGPPHETAAPKRTYDMMEGRVGRAISSASIEGLMGRAIPP-ERHSPHHLKEQHHIR 1352
                : |||   | |         |  | |  |   :|| | | || |   |
Db        615 EAGAQGPP-GPAGPAGE----RGEQGPA-GSPGFQGLPGPAGPPGEAGKPGE-------- 660

Qy       1353 GSITQGIPRSYVEAQEDYLRREAKLLKREGTPPPPPPSRDLTEAYKTQALGPLKLK--PA 1410
                   ||:|           | |        |   || |:             ||      |
Db        661 ----QGVPGDLGAPGPSGARGERGFPGERGVQGPPGPA-----------GPRGANGAPG 704

Qy       1411 HEGLVATVKEAGRSIHEIPREELRHTPELPLAPRPLKEGSITQGTP----LKYDTGASTT 1466
                 ::|     :||                  | ||     ||  || |    :  : ||:
Db         705 NDGAKG---DAG----------------APGAP-----GS--QGAPGLQGMPGERGAAGL 738

Qy       1467 GSKKHD-----VRSLIGSPGR------TFP--PVHPLDVMADARALERACYEESLKSRPG 1513
                  | |       :   ||||:      | |  |  |      |              |  |
Db         739 PGPKGDRGDAGPKGADGSPGKDGVRGLTGPIGPPGPAGAPGD------------KGESG 785

Qy       1514 TASSSGGSIARGAPVIVPELGKPRQSPLTYEDHGAPFAGHLPRGSPVTMREPTPRLQEGS 1573
                  :  :| : |||||     | | |    |  :   | | |      | |     ||      :|
Db         786 PSGPAGPTGARGAPGDRGEPGPP--GPAGFA--GPPGAD----GQPGAKGEPGDAGAKGD 837

Qy       1574 LSSSKASQDRKLTSTPREIAKSPHSTVPEHHPHPI----SPYEHLLRGVSGVDLYRSHIP 1629
                                  | |    |   | ||    :|      || :|
Db         838 ------------------AGPPGPAGPAGPPGPIGNVGAPGAKGARGSAGP-------- 870

Qy       1630 LAFDPTSIPRGIPLDAAAAYYLPRHLAPNPTYPHLYPPYLIRGYPDTAALENRQTIINDY 1689
                       |        ||    |    | |:  :  ||    | |  |  |  :    :
Db         871 --------PGATGFPGAAGRVGP----PGPS-GNAGPP----GPPGPAGKEGGKGPRGE- 912

Qy       1690 ITSQQMHHNTATAMAQRADML----------RGLSPRESSLALNYAAGPRGIIDLSQVP 1738
                        | | |   :           :|    :          ||:||      |
Db         913 ----------TGPAGRPGEVGPPGPPGPAGEKGSPGADGPAGAPGTPGPQGIAGQRGVV 961

Qy       1739 HLPVLVPPT--PGTPATAMDRLAYLPTAPQPFSSRHSSSPLSPGGPTHLTKPTTTSSSER 1796
                 ||       || |  : :       |    :     |  | ||  |  |   |  |
Db         962 GLPGQRGERGFPGLPGPSGE-----PGKQGPSGASGERGPPGPMGPPGLAGPPGESGRE- 1015

Qy       1797 ERDRDRERDRDRREREKSILTSTTTVEHAPIWRPGTEQSSGSSGSSGGGG-----GSSSRP 1851
                                           || | | |  || |  |      | :  |
Db        1016 ----------------------------GAPGAEGSPGRDGSPGAKGDRGETGPAGPP 1045

Qy       1852 ASHSHAHQHSPISPRTQDALQQRPSVLHNTGMKGIITAVEPSKPTVLRSTSTSSPVRPAA 1911
                  :     |: |              | |      |: |        : || |
Db        1046 GAPGAPGAPGPVGP--------------AGKSGDRGETGPAGP--------AGPVGPVG 1082

Qy       1912 TFPPATHCPLGGTLDGVYPTLMEPVLLPKEAPRVARPERPRADTGH-------------- 1957
                 |                              | |: || | |
Db        1083 ARGP----------------------------AGPQGPRGDKGETGEQGDRGIKGHRG 1112

Qy       1958 -AFLAKPPARSGLEPASSPSKGSEPRPLVPPVSGHATIARTPAKNLAPHHASPDPPAPPA 2016
                  : |  ||   |      ||  | |                              |  ||
Db        1113 FSGLQGPPGPPGSPGEQGPSGASGP--------------------------AGPRGPPG 1145

Qy       2017 SASDPHREKTQSKPFSIQELELRSLGYHGSSYSPEGVEPVSPVSSPSLTHDKGLPKHLEE 2076
                 ||  | ::      |  |       |  |          ||  |   |
Db        1146 SAGAPGKDGLNGLPGPI--------GPPGPRGRTGDAGPVGPPGPPG------------ 1184

Qy       2077 LDKSHLEGELRPKQPGPVKLGGEAAHLPHLRPLPESQPSSSPLLQTAPGVKGHQ------ 2130
                          |  |||    | : : ||         ||          |  | |
Db        1185 ----------PPGPPGPPSAGFDFSFLP--------QP---------PQEKAHDGGRYYR 1217

Qy       2131 ----------------RVVTLAQHISEVITQDYTRHHPQQLSAPLPAPLYSFPGASCPVL 2174
                             : :|:|  |   : : : :| :| :                 :|
Db        1218 ADDANVVRDRDLEVDTTLKSLSQQIENIRSPEGSRKNPAR---------------TC--R 1260

Qy       2175 DLRRPPSDLYLPPPDHGAPARGSPHSEGGKRSPEPNKTSVLGG---------GEDGIEPV 2225
                 ||:   ||                 : |:   :||:    |           ||  : |
Db        1261 DLKMCHSDW-----------------KSGEYWIDPNQGCNLDAIKVFCNMETGETCVYPT 1303

Qy       2226 SPPEGM------TEPGHSRSAVYPLLYRDGEQTEPSRMGSKSPGNTSQPPAFFSKLTESN 2279
                  |          |    :      || | |          |  | |     :||
Db        1304 QPSVAQKNWYISKNPKDKRHVWFGESMTDGFQFEYG-------GQGSDPADVAIQLT--- 1353

Qy       2280 SAMVKSKKQEINKKLNTHNRNEPEYNISQPG 2310
                 ::    | :: :  | :|    |   | |
Db        1354 --FLRLMSTEASQNITYHCKNSVAYMDQQTG 1382

RESULT 61
US-09-816-669A-14
; Sequence 14, Application US/09816669A
; Patent No. US20020137019A1
```

```
; GENERAL INFORMATION:
;  APPLICANT: GARABEDIAN, Michael
;  APPLICANT:  TANEJA, Samir
;  APPLICANT:  HITTELMAN, Adam
;  APPLICANT:  MARKUS, Steven
;  TITLE OF INVENTION: METHOD FOR SCREENING TRANSCRIPTIONAL COREGULATORY PROTEINS OF
;  TITLE OF INVENTION:  TRANSCRIPTION FACTORS,  AND ANDROGEN RECEPTOR TRANSCRIPTIONAL CO
;  TITLE OF INVENTION:  PROTEINS AS TARGETS FOR ANDROGEN RECEPTOR-DEPENDENT DISEASES
;  FILE REFERENCE: GARABEDIAN=1.1A
;  CURRENT APPLICATION NUMBER: US/09/816,669A
;  CURRENT FILING DATE:  2001-03-26
;  PRIOR APPLICATION NUMBER: 60/225,618
;  PRIOR FILING DATE: 2000-08-15
;  PRIOR APPLICATION NUMBER: 60/191,768
;  PRIOR FILING DATE: 2000-03-24
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
;   LENGTH: 2783
;   TYPE: PRT
;   ORGANISM: Human
US-09-816-669A-14

  Query Match             3.0%;  Score 394;  DB 9;  Length 2783;
  Best Local Similarity   18.2%;  Pred. No. 2.1e-09;
  Matches  510;  Conservative  273;  Mismatches  953;  Indels 1064;  Gaps  115;

Qy         99 EMEFIESKRPRLELLPDPLLRPSPLLATGQPAGSEDLT----KDRSLTGKLEPVSPPS-- 152
              |   :|     :: |   ||     ||| | |  :|| |    :|:     ||    |:
Db        535 ETSHLELSEADIQQLYGGLLANGDLLAMGDPTLAEDHTIIVEEDKEEESDLEDKQSPTGS 594

Qy        153 ---------------------PPHTDPELELVPPR-----LSKEELIQNMDRVDREITMV 186
                                  |:   |  | | |     : ||   |       : | :|
Db        595 DSGSVQEDSGSEPKRALPFRKGPNFTMEKFLDPSRPYKCTVCKESFTQ------KNILLV 648

Qy        187 E-QQISKLKKKQQQLEEEAAKPPEPEKPVSPPPIESKHRSLVQIIYDENRKKAEAAHRIL 245
               :| | | :: |:| |   |||       | :    ::
Db        649 HYNSVSHLHKLKRALQESATGQPEPTSSPDNKPFKCNTCNVA------------------ 690

Qy        246 EGLGPQVELPLYNQPSDTRQYHENIKINQAMRKKLILYFKRRNHARKQWKQKFCQRYDQL 305
                        |:| | | : |    ::|
Db        691 -----------YSQ-SSTLEIHMRSVLHQT---------------------------- 708

Qy        306 MEALEKKVERIENNPRRRAKESKVREYYEKQFPEIRKQRELQERMQSRVGQRGSGLSMSA 365
               :|   |:|    :       | :                |  | |    ::
Db        709 -KARAAKLEAASGSSNGTGNSSSI--------------SLSSSTPSPVSTSGSNTFTTS 752

Qy        366 ARSEHEVSEIIDGLSEQENLEKQMRQLAV-IPPMLYDADQQRIKFINMNGLMADPMKVYK 424
                 |         |::   ||  |:   :| :||                  : :|:
Db        753 NPSS-------AGIAPSSNLLSQVPTESVGMPP-----------------LGNPIGA-- 785

Qy        425 DRQVMNMWSEQE-KETFREKFMQHPKNFGLIASFLERKTVAECVLYYYLTKKNENYKSLV 483
                 |: |  | ||  |:|         :|||
Db        786 -----NIASPSEPKEANRKKLAD------MIAS--------------------------- 807

Qy        484 RRSYRRRGKSQQQQQQQQQQQQQQQQQPMPRSSQEEKD--EKEKEKEAEKEEEK------ 535
                 |: : |||||||||||||| | : ::  : :   ::| :::|   : :
Db        808 -----RQQQQQQQQQQQQQQQQQQQAQTLAQAQAQVQAHLQQELQQQAALIQSQLFNPTL 862

Qy        536 -PEVENDKEDLLKEKTD----------DTSGEDNDEKEAVASKGRKTANSQGRRKGRITR 584
                .|       | ||: :                : | |     : |  |    |   | :
Db        863 LPHFPMTTETLLQLQQQQHLLFPFYIPSAEFQLNPEVSLPVTSGALTLTGTG--PGLLED 920

Qy        585 SMAN---EANSEEAITPQQSAELAS--------------------MELNESSRWTEEEME 621
                 |        | : | |||     |                  : : |  : :: | :
Db        921 LKAQVQVPQQSHQQILPQQQQNQLSIAQSHSALLQPSQHPEKKNKLVIKEKEKESQRERD 980

Qy        622 TAKKGLLEHGRNWSAIARMVGSKTVSQCKNFYFNYKKRQNLDEILQQHKLKMEKERNARR 681
               :|: |    |                    |   :: | : |:                |:
Db        981 SAEGG---EG----------------------NTGPKETLPDALK----------AKE 1003

Qy        682 KKKKAPAAASEEAAPPPVVEDEEMEASGVSGN------------------EEEMVEEAE 722
               ||: ||   || :  || :      ||   ||                   :: |::

Db       1004 KKELAPGGGSEPSMLPPRI------ASDARGNATKALLENFGFELVIQYNENKQKVQKKN 1057

Qy        723 ALHASGNEVPRGECSGPATVNNS-----SDTESIPS---PHTEAAKDTGQNG-------P 767
                    |  : : ||      : ::     | | :     |  :  :    |        |
Db       1058 GKTDQGENLEKLECDSCGKLFSNILILKSHQEHVHQNYFPFKQLERFAKQYRDHYDKLYP 1117

Qy        768 KPPATLGADGPPPGPPTPPRRTSRAPIEP--TPASEATGAP--TPPPAPPSPSAP----- 818
               | |   : ||| || ||      || :|  |||  |:  |  :|  ||  || | |
Db       1118 LRPQT--PEPPPPPPPPPPPPLPAAPPQPASTPAIPASAPPITSPTIAPAQPSVPLTQLS 1175

Qy        819 -------------------------PPVV---------------------------- 822
                                       || :
Db       1176 MPMELPIFSPLMMQTMPLQTLPAQLPPQLGPVEPLPADLAQLYQHQLNPTLLQQQNKRPR 1235

Qy        823 --------------------PKEEKEEETAAAPPVEEGEEQK----------------- 844
                                  | |:| : | |     : |  ||
Db       1236 TRITDDQLRVLRQYFDINNSPSEEQIKEMAD----KSGLPQKVIKHWFRNTLFKERQRNK 1291

Qy        845 --------PP--AAEELAVDT-GKAEEPVKSEC-------------------------- 866
                     ||  : ||| :|:   : || | |
Db       1292 DSPYNFSNPPITSLEELKIDSRPPSPEPPKQEYWGSKRSSRTRFTDYQLRVLQDFFDANA 1351

Qy        867 ---TEEAE---------------------------EGPAKGKDAEAAEATAEGALKAEK 895
                :| |                           |   :||| |  | | :  ::
Db       1352 YPKDDEFEQLSNLLNLPTRVIVVWFQNARQKARKNYENQGEGKDGERRELTNDRYIRTSN 1411

Qy        896 KEGGSGRATTAKSSGAPQDSDSSATCSADEVDEAEGGDKNRLLSPRPSLLTPTGDPRANA 955
                : :                 |  || :| :   :|          :||||    :
Db       1412 LNYQCKKCSLVFQRIFDLIKHQKKLCYKDEDEEGQDDSQNEDSMDAMEILTPTSSSCSTP 1471

Qy        956 SPQKPLDL-KQLKQRAAAIPPIQVTKVHEP-----PREDAAPTKP----APPAPP---PP 1002
               | :            |:   :|:|    |        : :|    ||    || | |
Db       1472 MPSQAYSAPAPSANNTASSAFLQLTAEAEELATFNSKTEAGDEKPKLAEAPSAQPNQTQE 1531

Qy       1003 QNLQPESDAPQQPGSSPRGKSRSPAPPADKEAFAAEAQKLPGDPPCWTSGLPFPVPPREV 1062
               :  ||: :  ||     |  |: :|              ||||
Db       1532 KQGQPKPELQQQ--EQPEQKTNTP-----------QQKLP------------------- 1558

Qy       1063 IKASPHAPDPSAFSYAPPGHPLPLGLHDTARPVLPRPPTISNPP--PLISSAKHPSVLER 1120
                                       : | ||:||  : ||  ||  |:  || |
Db       1559 -------------------------QLVSLPSLPQPPPQAPPPQCPLPQSSPSPSQLS- 1591

Qy       1121 QIGAISQGMSVQLHVPYSE-HAKAPVGPVTMG---LPLPMDPKKLA--PFSGVKQEQ--- 1171
                           |:|    |    |      :   :|   |  |||    |   :: |
Db       1592 ------------HLPLKPLHTSTPQQLANLPPQLIPYQCDQCKLAFPSFEHWQEHQQLH 1638

Qy       1172 -LSPRGQAGPPE----SLGVP-------TAQEASVLRGTALGSVPGGSITKGIPSTRVPS 1219
                || : |   |:    || :|          || |   |:  :|  | |   |||  |:
Db       1639 FLSAQNQFIHPQFLDRSLDMPFMLFDPSNPLLASQLLSGAIPQIPASSATS--PST--PT 1694

Qy       1220 DSAITYRGSITHGTPADVLYKGTITRIIGEDSPSRLDRGREDSLPKG-HVIYEGKKGHVL 1278
                :  | :  :      |      |      | : | |  | | |    |:   ::|:    ::|
Db       1695 STMNTLKRKLEEKASAS---PGENDSGTGGEEPQRDKRLRTTITPEQLEILYQ---KYLL 1748

Qy       1279 SYEGGMSVTQCSKEDGRSSSGPPHETAAPKRTYDM------MEGRVGRAISSASIEGLMG 1332
                          |    :      ||    ||   :          | |:
Db       1749 D----------SNPTRKMLDHIAHEVGLKKRVVQVWFQNTRARERKGQF----------- 1787

Qy       1333 RAIPP---ERHSPH-------------HLKEQH-----------------------HIR 1352
                ||: |    |  |               |:: :|                       ::
Db       1788 RAVGPAQAHRRCPFCRALFKAKTALEAHIRSRHWHEAKRAGYNLTLSAMLLDCDGGLQMK 1847

Qy       1353 GSITQGIPRSYVEAQEDYLRREAKLLKREGTPPPPPPSRD-----LTEAYKTQALGPLKL 1407
                | |  |   |::                        || | |     |:   ||  | | |
Db       1848 GDIFDGTSFSHL----------------------PPSSSDGQGVPLSPVSKTMELSPRTL 1885

Qy       1408 KPAHEGLVATVKEAGRSIHEIPREELRHTPELPLAPRPLKEGSITQGTPLKYDTGASTTG 1467
                     :::|  |    | |      :  |       |      :      ||     |
Db       1886 LSP-----SSIKVEGIEDFESPS---MSSVNLNFDQTKLDNDDCSSVNTAITDTTTGDEG 1937

Qy       1468 SKKHDVRSLIGSPGRTFPPVHPLDVMADARALERACYEESLKSRPGTASSSGGSIARGAP 1527
                :  :|  |  |    |    | : :  |  :  : ||: | |  |  | | :
Db       1938 NADND--SATGIATETKSSSAPNEGLTKAAMMAMSEYEDRLSS--GLVSPA--------- 1984
```

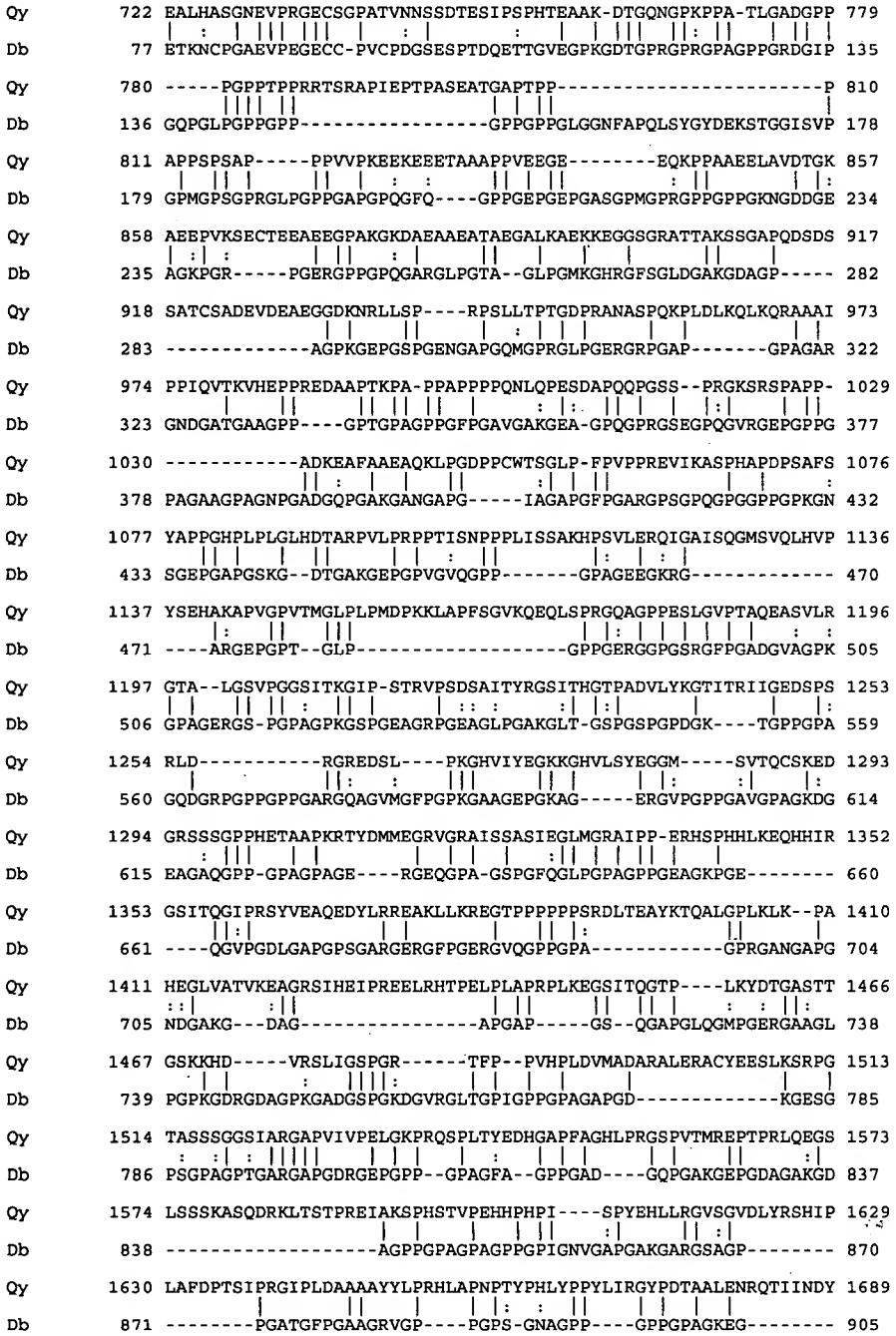
```
Qy       1528 VIVPELGKPRQSPLTYEDHGAPFAGHLPRGSPVTMREPTPRLQEGSLSSSKASQDR---- 1583
Db       1985 --------PSFYSKEYDNEGT--VDYSETSSLADPCSPSPG-ASGSAGKSGDSGDRPGQK 2033

Qy       1584 ----KLTSTPREIAKSPHSTVPEHHPHPISPYEHLLRGVSGVDLYRSHIPLAFDPTSIPR 1639
Db       2034 RFRTQMTNLQLKVLKSCFND----------------------YRTPTMLECEVLGNDI 2069

Qy       1640 GIPLDAAAAYY-------------LPRHLAPNPT-YPHLYPPYLIRGYPDTAALENRQTI 1685
Db       2070 GLPKRVVQVWFQNARAKEKKSKLSMAKHFGINQTSYEGPKTECTLCGIKYSARLSVRDHI 2129

Qy       1686 --------INDYITSQ----QMHHNTATA---MAQR-ADMLRGLSPRESSLALNYAAGPR 1729
Db       2130 FSQQHISKVKDTIGSQLDKEKEYFDPATVRQLMAQQELDRI-----KKANEVLGLAAQQQ 2184

Qy       1730 GIIDLS--QVPHLPVLVPPTPGTPATAMDRLAYLPTAPQPFSSRHSSSPLSPG---GPTH 1784
Db       2185 GMFDNTPLQALNLPTAYPALQGIPPVLLPGL--------------NSPSLPGFTPSNTA 2229

Qy       1785 LTKPTTTSSSERERDRDRERDRDRREREKSILTSTTTVEHAPIWRPGTEQSSGSSGSSGGG 1844
Db       2230 LTSPKPNLMG-------------------LPSTT------VPSPGL-PTSGLPNKPSSA 2262

Qy       1845 GGSSSRPASHSHAHQHSPISPRTQDALQQRPSVLHNTGMKGIITAVEPSKPTVLRSTSTS 1904
Db       2263 SLSSPTPAQATMA--MGPQQPPQQQQQQQQPQV------------QQPPPP--------- 2299

Qy       1905 SPVRPAATFPPATHCPLGGTLDGVYPTLMEPVLLPKEAPRVARPERPRADTGHAFLAKPP 1964
Db       2300 ----PAAQPPPTPQLPL----------QQQQQRKDKDSEKVKEKEKAHKGKGEP-LPVPK 2344

Qy       1965 ARSGLEPASSPSKGSEPRPLVPPVSGHATIARTPAKNLAPHHASPDPPAPPASASDPHRE 2024
Db       2345 KEKGEAPTATAATISAPLPTM-------EYAVDPAQLQA---------LQAALTSDPTAL 2388

Qy       2025 KTQSKPFSIQELELRSLGYHGSSYSPEGVEPVSPVSSPSLTHDKGLPKHLEELDKSHLEG 2084
Db       2389 LT-------------------SQFLPYFVPGFSPYYAPQI--------------------- 2409

Qy       2085 ELRPKQPGPVKLGGEAAHLPHLRPL-------PESQPSSSPLLQTAPG---VKGHQRVVT 2134
Db       2410 ------PGALQSG-------YLQPMYGMEGLFPYSPALSQALMGLSPGSLLQQYQQYQQS 2456

Qy       2135 LAQHISEVITQDYTRHHPQQLSAPLPAPLYSFPGASCPVLDLRRPPSDLYLPPPDHGAPA 2194
Db       2457 LQEAIQQQQQQKVQQQQPKASQTPVP------PGA----------------PSPDKD-PA 2493

Qy       2195 RGSPHSEGGKRSPE---------PNKTSVLGGGEDGI-EPVSPPE-----------GMT 2232
Db       2494 KESPKPEEQKNTPREVSPLLPKLPEEPEAESKSADSLYDPFIVPKVQYKLVCRKCQAGFS 2553

Qy       2233 EPGHSRSAVYPLLYRDGEQTEPSRM----------------GSKSPGNTSQPPAFFSKL 2275
Db       2554 DEEAARSHLKSLCFFGQSVVNLQEMVLHVPTGGGGGGSGGGGGGGGGGGGGGSYHCLA 2613

Qy       2276 TESNSAMVKSKKQEINKKLNTHN------RNEPEYNISQPGTEIFNMPAITGTGLMTYRS 2329
Db       2614 CESALCGEEALSQHLESALHKHRTITRAARNAKEHPSLLPHSACFPDPSTASTSQSAAHS 2673

Qy       2330 -----------QAVQEHASTNMGLEAIIRKALMGKYDQWEESPPLSANAFNPLNASASL 2377
Db       2674 NDSPPPPSAAAPSSASPHASRKSWPQVVSRAS--------AAKPP----SFPPLSSSSTV 2721

Qy       2378 PA----------AMPITAADGRSDHTLTSPGGGGKAKVSG 2407
Db       2722 TSSSCSTSGVQPSMPTDDYSEESDTDLSQKSDGPASPVEG 2761


RESULT 62
US-10-084-846A-5
; Sequence 5, Application US/10084846A
```

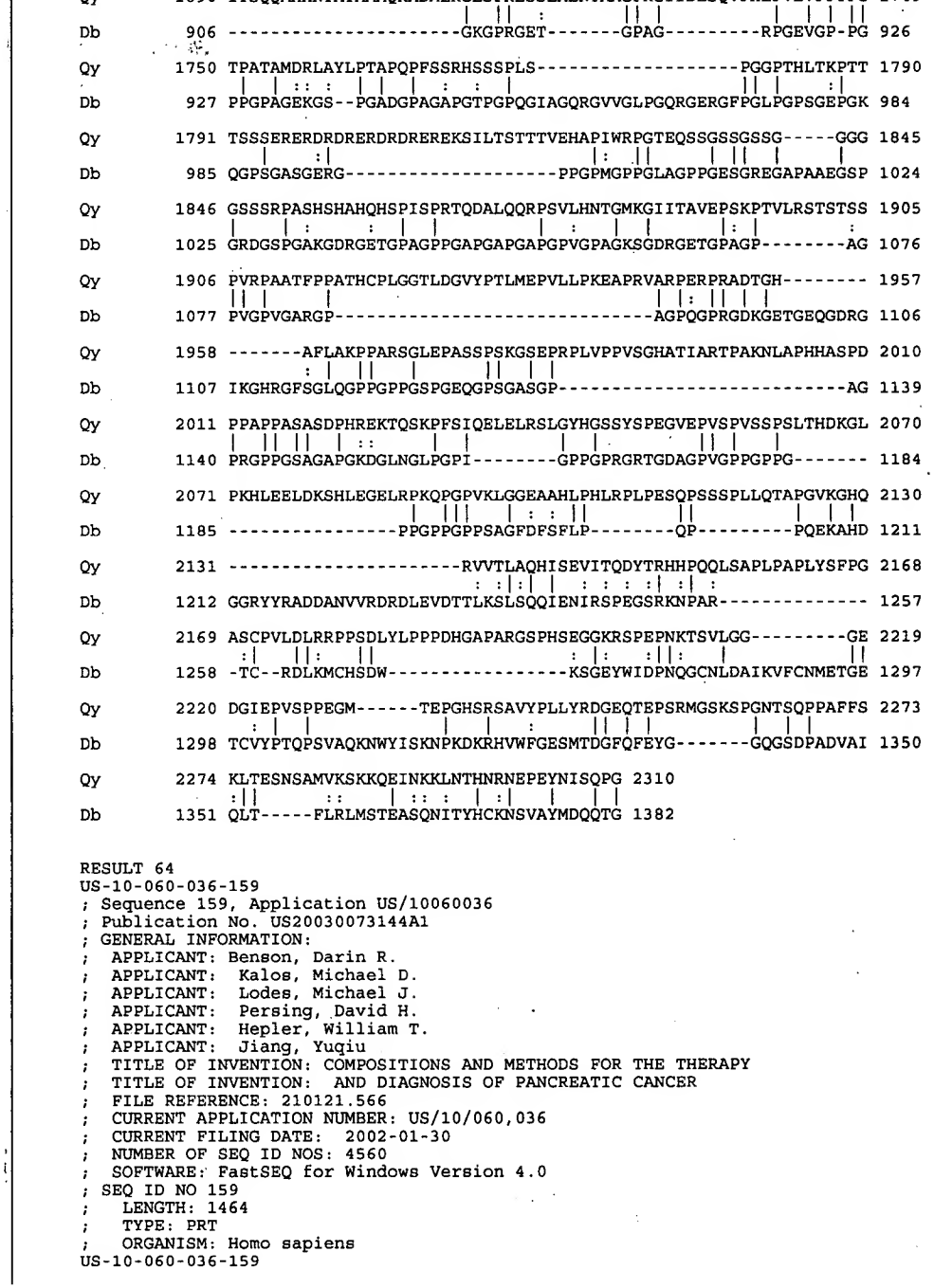
```
; Publication No. US20040006026A1
; GENERAL INFORMATION:
;  APPLICANT: WEITNAUER, GABRIELE
;  APPLICANT:  MUHLENWEG, AGNES
;  APPLICANT:  TREFZER, AXEL
;  APPLICANT:  BECHTHOLD, ANDREAS
;  TITLE OF INVENTION: AVILAMYCIN DERIVATIVES
;  FILE REFERENCE: 1974-005
;  CURRENT APPLICATION NUMBER: US/10/084,846A
;  CURRENT FILING DATE:  2003-02-25
;  PRIOR APPLICATION NUMBER: PCT/EP01/09815
;  PRIOR FILING DATE: 2001-08-24
;  PRIOR APPLICATION NUMBER: DE 101 09 166.4
;  PRIOR FILING DATE: 2001-02-25
;  NUMBER OF SEQ ID NOS: 120
;  SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 5
;   LENGTH: 19723
;   TYPE: PRT
;   ORGANISM: Streptomyces viridochromogenes
;   FEATURE:
;   OTHER INFORMATION: Protein 3: amino acid sequence encoded by coding strand 1.
;   OTHER INFORMATION: Start codon: atc, Start position: nucleotide 3.
US-10-084-846A-5

  Query Match             3.0%;  Score 394;  DB 15;  Length 19723;
  Best Local Similarity   20.2%;  Pred. No. 2.1e-08;
  Matches  481;  Conservative  168;  Mismatches  809;  Indels  924;  Gaps  114;

Qy        680 RRKKKKAPAAASEEAAFPPVVEDEEMEASGVSGNEEEMVEEAEAL---HASGNEVPRGEC 736
Db      17290 RRLVPAAPAPRGRAADRRPDGGRRPADRRGCTGQHRRTHRRGQARAPDHRTGRRAARAGG 17349

Qy        737 SGPATV-----NNSSDTESIPS---------PHTEAAKDTGQNG---------------- 766
Db      17350 GGPRGTPGRRGGRGADEEARPAAGRGGAGGRPAAHAVRGAGRTGTGAGGRGGDPRTDRAG 17409

Qy        767 -PKPPAT-----LGADGPPP----GPPTPPRRTSRAPIEPTPASEATGAPTPPP------ 810
Db      17410 RPRRCATAARPGLGRPAAPPGLRGGRPGRLRAALRRVLR--PARRALG-PRPRPRGGRGV 17466

Qy        811 -----------------APPS-----PSAPPPVVPKEEKEEETAAAPPVEEGEEQKPPAA 848
Db      17467 PAGRGGRAGQDPGARRHAPPGLTARLPGLRHPLRPRPAPDD-----PRRRDGARQDDPGH 17521

Qy        849 EELAVDTGKAEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGSGRATTAKS 908
Db      17522 RRTRAPSRRGAEPL------------------------HGGLSGQR------------- 17543

Qy        909 SGAPQDSDSSATCSADEVD----EAEGGDKNRLLSPRPS----LLTPTGDPRANASPQKP 960
Db      17544 -------------ADQLDPGDREAQCAARDTAARPRPAGRVRRLEGAGRRRGHH----- 17584

Qy        961 LDLKQLKQRAAAIPPIQ------VTKVHEPPREDAAPTKPAPP--------------AP 999
Db      17585 ------LRRAAGLPGARGRRGGHARRGRGPLREES----PDPPLHGRLRVGGALRPRAVP 17634

Qy       1000 PPPQNLQPESDAPQQPGSSPRGKSRSPAPPADKEAFAAEAQKLPGDPPCWTSGLPFPVPP 1059
Db      17635 DRYAHGEPRRGVP-QPGADP------PARAGGGHRRARRRGGLQGVPQGGRTGLSAPQP- 17686

Qy       1060 REVIKASPHAPDPSAFSYAPPGHPLPLGLHDTARPVLPRPPTISNPPPLISSAKHPSVLE 1119
Db      17687 ----AGRPHR-TPGA-----PAHGRVGGAQHGGRGGLPR----------------GRAR 17719

Qy       1120 RQIGAISQGMSVQLHVPYSEHAKAPVGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAG 1179
Db      17720 RQLHGDAQG------------GVRASGAVREGAPAARD-------RGGRRERAEGRGVLR 17760

Qy       1180 PPESLGVPTAQEASVLRGT-ALGSVPGGSITKGIPSTRVPSDSAITYRGSITHGTPADVL 1238
Db      17761 VPGRTGSGT-------RGAHGRGRARGGSRPAGRAGVR--------RGASHSGCAGPGG 17804
```

```
Qy      1239 YKGTITRIIGEDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKED-GRSS 1297
Db     17805 GKGSHA---GHAGPGG-DRRR-----TGHTGADGADG--AAQGAGTAGRACVRADLGQRA 17853

Qy      1298 SGPP-------HETAAPKRTYDMMEGRVG-------------RAISSASIEGLMGRAIPP 1337
Db     17854 ARPPTAARRRLRRRARSRRTARADRGRRSGPQHAGRLRGDPLRAAAQADRRTPGGRPCPP 17913

Qy      1338 ERHSPHHLKEQH----HIRGSITQGIPRSYVEAQEDYLRREAKLLKREGTPP----PPPP 1389
Db     17914 --HGPGPLGPRAPPALHGRGGRTPGADAG-----------------EQVPPVRRLRPPQ 17953

Qy      1390 SRDLTEAYKTQALGPLKLKPAHEG-----------------LVATVK------------ 1419
Db     17954 RGRRVDTGRRRHLGP-RPGPPHRGGGTGTVGGVSSGRRAYSVVAEVKARQAVGLTWRVGP 18012

Qy      1420 --------------EAGRSIH---EIPREELRHTPEL--------------PLAPRPLK 1447
Db     18013 SSGVFLESRMWMVWGQRATSTHWPPSLPLLDLNHSSAVLIYLPPDGLAGSSTPAESWPVT 18072

Qy      1448 EGSIT-----------------------------------QGTPLK-------YDTGASTT 1466
Db     18073 AASHTCAPGWTPRPGQKPTGCHVGSRCRPGAGDRASPGGRGRPVRRVRRCAWVETGGRAP 18132

Qy      1467 GSKKHDV----------------------RSL--------IGSPG-RTFP----PVHPLDV 1492
Db     18133 GVRRTRVGGGAGRGGLSAPRVRCTAPAPTRSSSSAGTCGAGSPGVRCDPPSWLPARPRRT 18192

Qy      1493 MADARALERACYEESLKSRPG--TASSSGGSIARGAPVIVPE---------LGKPRQSPL 1541
Db     18193 AVRARAGPRGRSRGSGAGRAGRRCRTSTGRWPPARAPCTAPRTAGTGRSPACSRPRQ-PT 18251

Qy      1542 TYEDHGA-PFAGHL---PRGSPVTMREPTPRLQEGSL---------------------- 1574
Db     18252 CPSCRGVRPAPGRVRAAGAASPVRSTRWSPR-RRGPLCPGNSASPPPCGRPASACPPSAW 18310

Qy      1575 ---SSSKA---------------------------------SQDRKLTSTPREIAKSPHS 1598
Db     18311 STGSSSAAASWSSCWPNSPHAPIRTRCRPVQNPSTGARTDKSQSRCATFVPTTQTNSPTS 18370

Qy      1599 TVPEHH------PHPISPYEHLLRGVSGVDLYRSHIPLAFDPTSIPRGIPLDAAAAYYLP 1652
Db     18371 ATSSTHATADPSPLPATPPGHSSTGAS---------PTGTGPTTY-------SSAASWRP 18414

Qy      1653 RHLAPNPTYPHLYPPYLIRGYPDTAALENRQTIINDYITSQQMHHNTATAMAQRADMLRG 1712
Db     18415 ---TPSST-------------------ASRPAMVSFFASAM-----TATSSASRSTTAER 18447

Qy      1713 LSPRESSLALNYAA----------------------GP----RGIIDLSQVPH------ 1739
Db     18448 AYPEPPTTPTKAAAVCSWSARWPTSGAWASERPAKSSGPSSPCRALVQFLWRPQSFGRTG 18507

Qy      1740 -----LPVLVPP-------TPGTP---ATAMDRLAYLPTAPQPFSSRHSSSPLSPGGPTH 1784
Db     18508 EAQVDMPLPTPPRRPHKRTTDGGPLRWQCVIPGEDWPPSARPPRPSPRASSPLSPPRP-- 18565

Qy      1785 LTKPTTTSSSERERDRDRERDRDREREKSILTST---TTVEHAPIWRPGTEQSSGSSGSS 1841
Db     18566 --PPRAVSST-----------RARQTPWPTATSTSRRTTPVRVP--RPDAPSSRSTAPTS 18610

Qy      1842 GG----GGGSSSRP----ASHSHAHQHSPISPRTQDALQQRPSVLHNTG------MKGII 1887
Db     18611 SARTRRPTATRSRPPRRRPRRSPPTRRSPPWSRTACSASTRPRPTRPPGAWTAWTRRTSP 18670

Qy      1888 TAVEP------------SKPTVLRSTSTSSPVRPAATFPPATHCPLGGTLDGVYPTLMEP 1935
Db     18671 TARTPTRTRPGRERRTSSTPASASPTATSAAGPPTATTPSTTTTP--------------- 18715

Qy      1936 VLLPKEAPRVAR-------------PERPRADTGHAFLAK-PPARSGLEPASSPSKG--- 1978
Db     18716 ---PRTATATARTWPARSPATPTASPRRPRSASACTTPARAPPPRS--SPASTGSPGTPS 18770

Qy      1979 SEPRPLVPPVSGHATIARTPAKNLAPHHASPDPPAPPASASDPHREKTQSKPFSIQELEL 2038
```

```
Db     18771 SRPSPTCPS-------AAAPTRPSTRPYATPWPPASPS----PWRPAT-SRP-------- 18810

Qy      2039 RSLGYHGSSYSPEGVEPVSPVSSPSLTHDKGLPKHLEELDKSHLEGELRPKQPGPVKLGG 2098
Db     18811 ---------TPPRGHPHASPRPSRSARRPARTP--------------RPATP---TTAP 18843

Qy      2099 EAAHLPHLRPLPE------SQPSSSPLLQTAPGVKGHQRVVTLAQHISEVITQDYTRHHP 2152
Db     18844 SSTSSPPVRPSPRPGTQATRRPTPSPVRRWRPRT--------------------WRAPP 18882

Qy      2153 QQLSAPLPAPLYSFPGASCPVLDLRRPPSDLYLPPPDHGAPAR----GSPHSEGG----- 2203
Db     18883 RSTSPPTPRP----PRPRSP---RRRPPP----PPASSPTPARARPTGSCTSAGGTTTLR 18931

Qy      2204 ---KRSPEPNKTSVLGGGEDGIEPVSPPEGMTEP-------------GHSRSAVYPLLYR 2247
Db     18932 ARASRTPVTTRSATTPRSSPRRSPASP---ATRPRPSPRSTSSTRTSATSRSSSPPTARR 18988

Qy      2248 D-GEQTEPSRMGSKS-------------------------PGNT-----SQPPAFFSKL 2275
Db     18989 TRSSRTAPAAVRTTSTPRTRTPPRRRPTARGNCGATTRTTTPGATPGRCSSSPPPVLVVL 19048

Qy      2276 TESNSAMVKSKKQEINKKLNTHNRNEPEYNISQPGTEIFNMPAITGTGLMTYRSQAVQEH 2335
Db     19049 AARHRVPTRPTAR------TTHVR--PRHAASGKGVEARWAGWVRGASPQALRAGGVPSY 19100

Qy      2336 ASTNMG----------------------LEAIIRKALMGKYDQWEE-----------S 2360
Db     19101 ACCADGARTPTGPQSRTAGRGRVGGCPPAVTGVNTRHLLAQGCRRSEDGHPPARPRPTTQ 19160

Qy      2361 PPLSANAFNPLNASASLPAAMPITAADGRSDHTLTSPGGGGKAKVSGRPSSRK-AKSPAP 2419
Db     19161 PKATRTPTHPPQAAAGMRTRTPTTAG---------SPPSGSPFRPASPPAYHSPSRTPVP 19211

Qy      2420 GLASGDRPPSV--SSVHSEGDCNRRTPLTNRVWEDRPSSAGSTPFPYNPLIMRLQA---G 2474
Db     19212 A-ARPCRPNQEHGTRVHTSAP---RTPTAARVPGAAPRAERAGPVPATVAIPARRTRGPG 19267

Qy      2475 VMASPPPPGLPAGSGP----------LAGPHHAWDEEPKPLL 2506
Db     19268 TVRVPPPLRLPQPLRPPRTRQRRRHRRPGPRH-------PLL 19302
```

```
RESULT 63
US-09-918-715-261
; Sequence 261, Application US/09918715
; Publication No. US20030017157A1
; GENERAL INFORMATION:
;  APPLICANT: Brad St. Croix
;  APPLICANT:  Bert Vogelstein
;  APPLICANT:  Kenneth Kinzler
;  TITLE OF INVENTION: ENDOTHELIAL CELL EXPRESSION PATTERNS
;  FILE REFERENCE: 1107.00134
;  CURRENT APPLICATION NUMBER: US/09/918,715
;  CURRENT FILING DATE:  2001-08-01
;  PRIOR APPLICATION NUMBER: 60/222,599
;  PRIOR FILING DATE: 2000-08-02
;  PRIOR APPLICATION NUMBER: 60/224,360
;  PRIOR FILING DATE: 2000-08-11
;  PRIOR APPLICATION NUMBER: 60/282,850
;  PRIOR FILING DATE: 2000-04-11
;  NUMBER OF SEQ ID NOS: 358
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 261
;   LENGTH: 1464
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-918-715-261

  Query Match             3.0%;  Score 392.5;  DB 12;  Length 1464;
  Best Local Similarity   21.3%;  Pred. No. 1.1e-09;
  Matches  378;  Conservative  102;  Mismatches  638;  Indels  659;  Gaps   84;
```





RESULT 64
US-10-060-036-159
; Sequence 159, Application US/10060036
; Publication No. US20030073144A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Persing, David H.
; APPLICANT: Hepler, William T.
; APPLICANT: Jiang, Yuqiu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF PANCREATIC CANCER
; FILE REFERENCE: 210121.566
; CURRENT APPLICATION NUMBER: US/10/060,036
; CURRENT FILING DATE: 2002-01-30
; NUMBER OF SEQ ID NOS: 4560
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 159
; LENGTH: 1464
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-060-036-159

```
Query Match             3.0%;  Score 392.5;  DB 14;  Length 1464;
Best Local Similarity   21.3%;  Pred. No. 1.1e-09;
Matches  378;  Conservative  102;  Mismatches  638;  Indels  659;  Gaps   84;

Qy        722 EALHASGNEVPRGECSGPATVNNSSDTESIPSPHTEAAK-DTGQNGPKPPA-TLGADGPP 779
              |  :  | ||| |||  |   : |       :   |  | |||  ||: ||   | || |
Db         77 ETKNCPGAEVPEGECC-PVCPDGSESPTDQETTGVEGPKGDTGPRGPRGPAGPPGRDGIP 135

Qy        780 -----PGPPTPPRRTSRAPIEPTPASEATGAPTPP-----------------------P 810
                  |||| ||                  | | ||                      |
Db        136 GQPGLPGPPGPP----------------GPPGPPGLGGNFAPQLSYGYDEKSTGGISVP 178

Qy        811 APPSPSAP-----PPVVPKEEKEEETAAAPPVEEGE--------EQKPPAAEELAVDTGK 857
              |  || |     ||  |  :  :     || | ||          : ||       | |:
Db        179 GPMGPSGPRGLPGPPGAPGPQGFQ----GPPGEPGEPGASGPMGPRGPPGPPGKNGDDGE 234

Qy        858 AEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGSGRATTAKSSGAPQDSDS 917
              | :| :       | ||   : |      ||   |   |       ||    |
Db        235 AGKPGR-----PGERGPPGPQGARGLPGTA--GLPGMKGHRGFSGLDGAKGDAGP----- 282

Qy        918 SATCSADEVDEAEGGDKNRLLSP----RPSLLTPTGDPRANASPQKPLDLKQLKQRAAAI 973
                      | |    11     |  : | | |     |  |          | |
Db        283 ------------AGPKGEPGSPGENGAPGQMGPRGLPGERGRPGAP-------GPAGAR 322

Qy        974 PPIQVTKVHEPPREDAAPTKPA-PPAPPPPQNLQPESDAPQQPGSS--PRGKSRSPAPP- 1029
                  |    ||      || || ||  |      : |:  || |   |   |:|    | ||
Db        323 GNDGATGAAGPP----GPTGPAGPPGFPGAVGAKGEA-GPQGPRGSEGPQGVRGEPGPPG 377

Qy       1030 ------------ADKEAFAAEAQKLPGDPPCWTSGLP-FPVPPREVIKASPHAPDPSAFS 1076
                        || :  |   |   ||      :| | ||          | |     :
Db        378 PAGAAGPAGNPGADGQPGAKGANGAPG-----IAGAPGFPGARGPSGPQGPGGPPGPKGN 432

Qy       1077 YAPPGHPLPLGLHDTARPVLPRPPTISNPPPLISSAKHPSVLERQIGAISQGMSVQLHVP 1136
              || |   |  ||    | |  :  ||          |:  | : |
Db        433 SGEPGAPGSKG--DTGAKGEPGPVGVQGPP-------GPAGEEGKRG------------- 470

Qy       1137 YSEHAKAPVGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPESLGVPTAQEASVLR 1196
                |:   ||  |||                  | |: | | | | | |   :  :
Db        471 ----ARGEPGPT--GLP------------------GPPGERGGPGSRGFPGADGVAGPK 505

Qy       1197 GTA--LGSVPGGSITKGIP-STRVPSDSAITYRGSITHGTPADVLYKGTITRIIGEDSPS 1253
              | |   || || :  || |     | :: :     :| |:|       |       |   |:
Db        506 GPAGERGS-PGPAGPKGSPGEAGRPGEAGLPGAKGLT-GSPGSPGPDGK----TGPPGPA 559

Qy       1254 RLD-----------RGREDSL----PKGHVIYEGKKGHVLSYEGGM-----SVTQCSKED 1293
              |           ||:  :   |||     || |     | |:   :|   |:
Db        560 GQDGRPGPPGPPGARGQAGVMGFPGPKGAAGEPGKAG-----ERGVPGPPGAVGPAGKDG 614

Qy       1294 GRSSSGPPHETAAPKRTYDMMEGRVGRAISSASIEGLMGRAIPP-ERHSPHHLKEQHHIR 1352
                : |||   | |          | | |  |   :|| | | || |   |
Db        615 EAGAQGPP-GPAGPAGE----RGEQGPA-GSPGFQGLPGPAGPPGEAGKPGE-------- 660

Qy       1353 GSITQGIPRSYVEAQEDYLRREAKLLKREGTPPPPPPSRDLTEAYKTQALGPLKLK--PA 1410
              ||:|           | |       |   || |:           ||      |
Db        661 ----QGVPGDLGAPGPSGARGERGFPGERGVQGPPGPA-----------GPRGANGAPG 704

Qy       1411 HEGLVATVKEAGRSIHEIPREELRHTPELPLAPRPLKEGSITQGTP----LKYDTGASTT 1466
              ::|      :||               | ||     ||  || |   :  : ||:
Db        705 NDGAKG---DAG---------------APGAP-----GS--QGAPGLQGMPGERGAAGL 738

Qy       1467 GSKKHD-----VRSLIGSPGR------TFP--PVHPLDVMADARALERACYEESLKSRPG 1513
                | |      :   ||||:       | |  |  |      |               |   |
Db        739 PGPKGDRGDAGPKGADGSPGKDGVRGLTGPIGPPGPAGAPGD------------KGESG 785

Qy       1514 TASSSGGSIARGAPVIVPELGKPRQSPLTYEDHGAPFAGHLPRGSPVTMREPTPRLQEGS 1573
                :  :| : |||||     | | |    |  :   |.| |  :    | |     ||     :|
Db        786 PSGPAGPTGARGAPGDRGEPGPP--GPAGFA--GPPGAD----GQPGAKGEPGDAGAKGD 837

Qy       1574 LSSSKASQDRKLTSTPREIAKSPHSTVPEHHPHPI----SPYEHLLRGVSGVDLYRSHIP 1629
                                 | |   |  | |  |  :|     || :|
Db        838 ------------------AGPPGPAGPAGPPGPIGNVGAPGAKGARGSAGP-------- 870

Qy       1630 LAFDPTSIPRGIPLDAAAAYYLPRHLAPNPTYPHLYPPYLIRGYPDTAALENRQTIINDY 1689
                                                                          |
Db                                                                        |
                                                                          |
Db        871 --------PGATGFPGAAGRVGP----PGPS-GNAGPP----GPPGPAGKEG-------- 905

Qy       1690 ITSQQMHHNTATAMAQRADMLRGLSPRESSLALNYAAGPRGIIDLSQVPHLPVLVPPTPG 1749
                                       |  ||  :       || |          |  | | ||
Db        906 ----------------------GKGPRGET-------GPAG---------RPGEVGP-PG 926

Qy       1750 TPATAMDRLAYLPTAPQPFSSRHSSSPLS-------------------PGGPTHLTKPTT 1790
                |  | :: :  | |  |  :  :  |                     || |    :|
Db        927 PPGPAGEKGS--PGADGPAGAPGTPGPQGIAGQRGVVGLPGQRGERGFPGLPGPSGEPGK 984

Qy       1791 TSSSERERDRDRERDRDRERERKSILTSTTTVEHAPIWRPGTEQSSGSSGSSG-----GGG 1845
                  |    :|                        |:  ||      | ||  |      |
Db        985 QGPSGASGERG-------------------PPGPMGPPGLAGPPGESGREGAPAAEGSP 1024

Qy       1846 GSSSRPASHSHAHQHSPISPRTQDALQQRPSVLHNTGMKGIITAVEPSKPTVLRSTSTSS 1905
                |     | :     :  | |                 | :   | |       |: |        :
Db       1025 GRDGSPGAKGDRGETGPAGPPGAPGAPGAPGPVGPAGKSGDRGETGPAGP--------AG 1076

Qy       1906 PVRPAATFPPATHCPLGGTLDGVYPTLMEPVLLPKEAPRVARPERPRADTGH-------- 1957
              || |      |                                 | |: || | |
Db       1077 PVGPVGARGP-----------------------------AGPQGPRGDKGETGEQGDRG 1106

Qy       1958 -------AFLAKPPARSGLEPASSPSKGSEPRPLVPPVSGHATIARTPAKNLAPHHASPD 2010
                    : |  ||    |       ||  | |
Db       1107 IKGHRGFSGLQGPPGPPGSPGEQGPSGASGP-------------------------AG 1139

Qy       2011 PPAPPASASDPHREKTQSKPFSIQELELRSLGYHGSSYSPEGVEPVSPVSSPSLTHDKGL 2070
              |  || ||  | ::      |  |         |  |             ||  |    |
Db       1140 PRGPPGSAGAPGKDGLNGLPGPI--------GPPGPRGRTGDAGPVGPPGPPG------- 1184

Qy       2071 PKHLEELDKSHLEGELRPKQPGPVKLGGEAAHLPHLRPLPESQPSSSPLLQTAPGVKGHQ 2130
                                 |  |||    | : : ||         ||          | | |
Db       1185 ----------------PPGPPGPPSAGFDFSFLP--------QP---------PQEKAHD 1211

Qy       2131 ----------------------RVVTLAQHISEVITQDYTRHHPQQLSAPLPAPLYSFPG 2168
                                    : :|:| |  : : : :| :| :
Db       1212 GGRYYRADDANVVRDRDLEVDTTLKSLSQQIENIRSPEGSRKNPAR-------------- 1257

Qy       2169 ASCPVLDLRRPPSDLYLPPPDHGAPARGSPHSEGGKRSPEPNKTSVLGG---------GE 2219
                :|   ||:   ||                 : |:   :||:   |           ||
Db       1258 -TC--RDLKMCHSDW----------------KSGEYWIDPNQGCNLDAIKVFCNMETGE 1297

Qy       2220 DGIEPVSPPEGM------TEPGHSRSAVYPLLYRDGEQTEPSRMGSKSPGNTSQPPAFFS 2273
                :| |              |   |   :     || | |           | | |
Db       1298 TCVYPTQPSVAQKNWYISKNPKDKRHVWFGESMTDGFQFEYG-------GQGSDPADVAI 1350

Qy       2274 KLTESNSAMVKSKKQEINKKLNTHNRNEPEYNISQPG 2310
              :||      ::    | :: :  | :|   |   | |
Db       1351 QLT-----FLRLMSTEASQNITYHCKNSVAYMDQQTG 1382
```

RESULT 65
US-10-171-311-36
; Sequence 36, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatt, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoersh, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13

```
;  PRIOR APPLICATION NUMBER: US 60/298,155
;  PRIOR FILING DATE: 2001-06-13
;  PRIOR APPLICATION NUMBER: US 60/335,936
;  PRIOR FILING DATE: 2001-11-14
;  NUMBER OF SEQ ID NOS: 238
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 36
;   LENGTH: 1464
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-171-311-36

  Query Match             3.0%;  Score 392.5;  DB 14;  Length 1464;
  Best Local Similarity   21.3%;  Pred. No. 1.1e-09;
  Matches  378;  Conservative  102;  Mismatches  638;  Indels  659;  Gaps   84;

Qy       722 EALHASGNEVPRGECSGPATVNNSSDTESIPSPHTEAAK-DTGQNGPKPPA-TLGADGPP 779
Db        77 ETKNCPGAEVPEGECC-PVCPDGSESPTDQETTGVEGPKGDTGPRGPRGPAGPPGRDGIP 135

Qy       780 -----PGPPTPPRRTSRAPIEPTPASEATGAPTPP-----------------------P 810
Db       136 GQPGLPGPPGPP----------------GPPGPPGLGGNFAPQLSYGYDEKSTGGISVP 178

Qy       811 APPSPSAP-----PPVVPKEEKEEETAAAPPVEEGE--------EQKPPAAEELAVDTGK 857
Db       179 GPMGPSGPRGLPGPPGAPGPQGFQ----GPPGEPGEPGASGPMGPRGPPGPPGKNGDDGE 234

Qy       858 AEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGSGRATTAKSSGAPQDSDS 917
Db       235 AGKPGR-----PGERGPPGPQGARGLPGTA--GLPGMKGHRGFSGLDGAKGDAGP----- 282

Qy       918 SATCSADEVDEAEGGDKNRLLSP----RPSLLTPTGDPRANASPQKPLDLKQLKQRAAAI 973
Db       283 ------------AGPKGEPGSPGENGAPGQMGPRGLPGERGRPGAP-------GPAGAR 322

Qy       974 PPIQVTKVHEPPREDAAPTKPA-PPAPPPPQNLQPESDAPQQPGSS--PRGKSRSPAPP- 1029
Db       323 GNDGATGAAGPP----GPTGPAGPPGFPGAVGAKGEA-GPQGPRGSEGPQGVRGEPGPPG 377

Qy      1030 ------------ADKEAFAAEAQKLPGDPPCWTSGLP-FPVPPREVIKASPHAPDPSAFS 1076
Db       378 PAGAAGPAGNPGADGQPGAKGANGAPG-----IAGAPGFPGARGPSGPQGPGGPPGPKGN 432

Qy      1077 YAPPGHPLPLGLHDTARPVLPRPPTISNPPPLISSAKHPSVLERQIGAISQGMSVQLHVP 1136
Db       433 SGEPGAPGSKG--DTGAKGEPGPVGVQGPP-------GPAGEEGKRG------------- 470

Qy      1137 YSEHAKAPVGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPESLGVPTAQEASVLR 1196
Db       471 ----ARGEPGPT--GLP------------------GPPGERGGPGSRGFPGADGVAGPK 505

Qy      1197 GTA--LGSVPGGSITKGIP-STRVPSDSAITYRGSITHGTPADVLYKGTITRIIGEDSPS 1253
Db       506 GPAGERGS-PGPAGPKGSPGEAGRPGEAGLPGAKGLT-GSPGSPGPDGK----TGPPGPA 559

Qy      1254 RLD-----------RGREDSL----PKGHVIYEGKKGHVLSYEGGM-----SVTQCSKED 1293
Db       560 GQDGRPGPPGPPGARGQAGVMGFPGPKGAAGEPGKAG-----ERGVPGPPGAVGPAGKDG 614

Qy      1294 GRSSSGPPHETAAPKRTYDMMEGRVGRAISSASIEGLMGRAIPP-ERHSPHHLKEQHHIR 1352
Db       615 EAGAQGPP-GPAGPAGE----RGEQGPA-GSPGFQGLPGPAGPPGEAGKPGE-------- 660

Qy      1353 GSITQGIPRSYVEAQEDYLRREAKLLKREGTPPPPPPSRDLTEAYKTQALGPLKLK--PA 1410
Db       661 ----QGVPGDLGAPGPSGARGERGFPGERGVQGPPGPA-----------GPRGANGAPG 704

Qy      1411 HEGLVATVKEAGRSIHEIPREELRHTPELPLAPRPLKEGSITQGTP----LKYDTGASTT 1466
Db       705 NDGAKG---DAG---------------APGAP-----GS--QGAPGLQGMPGERGAAGL 738

Qy      1467 GSKKHD-----VRSLIGSPGR------TFP--PVHPLDVMADARALERACYEESLKSRPG 1513
```

```
Db       739 PGPKGDRGDAGPKGADGSPGKDGVRGLTGPIGPPGPAGAPGD------------KGESG 785

Qy      1514 TASSSGGSIARGAPVIVPELGKPRQSPLTYEDHGAPFAGHLPRGSPVTMREPTPRLQEGS 1573
Db       786 PSGPAGPTGARGAPGDRGEPGPP--GPAGFA--GPPGAD----GQPGAKGEPGDAGAKGD 837

Qy      1574 LSSSKASQDRKLTSTPREIAKSPHSTVPEHHPHPI----SPYEHLLRGVSGVDLYRSHIP 1629
Db       838 ------------------AGPPGPAGPAGPPGPIGNVGAPGAKGARGSAGP-------- 870

Qy      1630 LAFDPTSIPRGIPLDAAAAYYLPRHLAPNPTYPHLYPPYLIRGYPDTAALENRQTIINDY 1689
Db       871 --------PGATGFPGAAGRVGP----PGPS-GNAGPP----GPPGPAGKEG-------- 905

Qy      1690 ITSQQMHHNTATAMAQRADMLRGLSPRESSLALNYAAGPRGIIDLSQVPHLPVLVPPTPG 1749
Db       906 ----------------------GKGPRGET-------GPAG---------RPGEVGP-PG 926

Qy      1750 TPATAMDRLAYLPTAPQPFSSRHSSSPLS------------------PGGPTHLTKPTT 1790
Db       927 PPGPAGEKGS--PGADGPAGAPGTPGPQGIAGQRGVVGLPGQRGERGFPGLPGPSGEPGK 984

Qy      1791 TSSSERERDRDRERDRDRERERKSILTSTTTVEHAPIWRPGTEQSSGSSGSSG-----GGG 1845
Db       985 QGPSGASGERG-------------------PPGPMGPPGLAGPPGESGREGAPAAEGSP 1024

Qy      1846 GSSSRPASHSHAHQHSPISPRTQDALQQRPSVLHNTGMKGIITAVEPSKPTVLRSTSTSS 1905
Db      1025 GRDGSPGAKGDRGETGPAGPPGAPGAPGAPGPVGPAGKSGDRGETGPAGP--------AG 1076

Qy      1906 PVRPAATFPPATHCPLGGTLDGVYPTLMEPVLLPKEAPRVARPERPRADTGH-------- 1957
Db      1077 PVGPVGARGP----------------------------AGPQGPRGDKGETGEQGDRG 1106

Qy      1958 -------AFLAKPPARSGLEPASSPSKGSEPRPLVPPVSGHATIARTPAKNLAPHHASPD 2010
Db      1107 IKGHRGFSGLQGPPGPPGSPGEQGPSGASGP--------------------------AG 1139

Qy      2011 PPAPPASASDPHREKTQSKPFSIQELELRSLGYHGSSYSPEGVEPVSPVSSPSLTHDKGL 2070
Db      1140 PRGPPGSAGAPGKDGLNGLPGPI--------GPPGPRGRTGDAGPVGPPGPPG------- 1184

Qy      2071 PKHLEELDKSHLEGELRPKQPGPVKLGGEAAHLPHLRPLPESQPSSSPLLQTAPGVKGHQ 2130
Db      1185 ----------------PPGPPGPPSAGFDFSFLP--------QP---------PQEKAHD 1211

Qy      2131 ----------------------RVVTLAQHISEVITQDYTRHHPQQLSAPLPAPLYSFPG 2168
Db      1212 GGRYYRADDANVVRDRDLEVDTTLKSLSQQIENIRSPEGSRKNPAR------------- 1257

Qy      2169 ASCPVLDLRRPPSDLYLPPPDHGAPARGSPHSEGGKRSPEPNKTSVLGG---------GE 2219
Db      1258 -TC--RDLKMCHSDW----------------KSGEYWIDPNQGCNLDAIKVFCNMETGE 1297

Qy      2220 DGIEPVSPPEGM------TEPGHSRSAVYPLLYRDGEQTEPSRMGSKSPGNTSQPPAFFS 2273
Db      1298 TCVYPTQPSVAQKNWYISKNPKDKRHVWFGESMTDGFQFEYG-------GQGSDPADVAI 1350

Qy      2274 KLTESNSAMVKSKKQEINKKLNTHNRNEPEYNISQPG 2310
Db      1351 QLT-----FLRLMSTEASQNITYHCKNSVAYMDQQTG 1382

RESULT 66
US-10-149-352-2
; Sequence 2, Application US/10149352
; Publication No. US20030105050A1
; GENERAL INFORMATION:
;  APPLICANT: Beri, Rajinder
;  TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDES
;  FILE REFERENCE: 06275-254US1
;  CURRENT APPLICATION NUMBER: US/10/149,352
```

```
;  CURRENT FILING DATE:  2002-06-10
;  PRIOR APPLICATION NUMBER: PCT/GB00/04741
;  PRIOR FILING DATE: 2000-12-12
;  PRIOR APPLICATION NUMBER: GB 9929487.8
;  PRIOR FILING DATE: 1999-12-15
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: PatentIn Ver. 4.0
; SEQ ID NO 2
;   LENGTH: 1464
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-149-352-2

  Query Match             3.0%;  Score 392.5;  DB 14;  Length 1464;
  Best Local Similarity   21.3%;  Pred. No. 1.1e-09;
  Matches  378;  Conservative  102;  Mismatches  638;  Indels  659;  Gaps   84;

Qy        722 EALHASGNEVPRGECSGPATVNNSSDTESIPSPHTEAAK-DTGQNGPKPPA-TLGADGPP 779
              |  :  | ||| |||  |  : |     :  | | |||  ||: ||  | || |
Db         77 ETKNCPGAEVPEGECC-PVCPDGSESPTDQETTGVEGPKGDTGPRGPRGPAGPPGRDGIP 135

Qy        780 -----PGPPTPPRRTSRAPIEPTPASEATGAPTPP----------------------P 810
                  |||| ||                 | | ||                     |
Db        136 GQPGLPGPPGPP----------------GPPGPPGLGGNFAPQLSYGYDEKSTGGISVP 178

Qy        811 APPSPSAP-----PPVVPKEEKEEETAAAPPVEEGE--------EQKPPAAEELAVDTGK 857
               |  || |      ||  |  :  :     || | ||         : ||       | |:
Db        179 GPMGPSGPRGLPGPPGAPGPQGFQ----GPPGEPGEPGASGPMGPRGPPGPPGKNGDDGE 234

Qy        858 AEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGSGRATTAKSSGAPQDSDS 917
              | :| :       | ||   : |    ||   |    |    |      ||    |
Db        235 AGKPGR-----PGERGPPGPQGARGLPGTA--GLPGMKGHRGFSGLDGAKGDAGP----- 282

Qy        918 SATCSADEVDEAEGGDKNRLLSP----RPSLLTPTGDPRANASPQKPLDLKQLKQRAAAI 973
                        | |    ||     | : | | |     | |          | |
Db        283 ------------AGPKGEPGSPGENGAPGQMGPRGLPGERGRPGAP-------GPAGAR 322

Qy        974 PPIQVTKVHEPPREDAAPTKPA-PPAPPPPQNLQPESDAPQQPGSS--PRGKSRSPAPP- 1029
                  |    ||     || || ||  |     : |:  || |  |  |:|    | ||
Db        323 GNDGATGAAGPP----GPTGPAGPPGFPGAVGAKGEA-GPQGPRGSEGPQGVRGEPGPPG 377

Qy       1030 ------------ADKEAFAAEAQKLPGDPPCWTSGLP-FPVPPREVIKASPHAPDPSAFS 1076
                         ||  :  |  |  ||      :| | ||          |  |      :
Db        378 PAGAAGPAGNPGADGQPGAKGANGAPG-----IAGAPGFPGARGPSGPQGPGGPPGPKGN 432

Qy       1077 YAPPGHPLPLGLHDTARPVLPRPPTISNPPPLISSAKHPSVLERQIGAISQGMSVQLHVP 1136
                |  |  |  | ||     | |  :  ||        |:  | : |
Db        433 SGEPGAPGSKG--DTGAKGEPGPVGVQGPP-------GPAGEEGKRG------------- 470

Qy       1137 YSEHAKAPVGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPESLGVPTAQEASVLR 1196
                 |:   ||   |||                    | |: | | | | | |   :  :
Db        471 ----ARGEPGPT--GLP------------------GPPGERGGPGSRGFPGADGVAGPK 505

Qy       1197 GTA--LGSVPGGSITKGIP-STRVPSDSAITYRGSITHGTPADVLYKGTITRIIGEDSPS 1253
              | |   || || :  || |     | :: :     :| |:|      |     |    |:
Db        506 GPAGERGS-PGPAGPKGSPGEAGRPGEAGLPGAKGLT-GSPGSPGPDGK----TGPPGPA 559

Qy       1254 RLD-----------RGREDSL----PKGHVIYEGKKGHVLSYEGGM-----SVTQCSKED 1293
                |          ||:  :    |||     || |     | |:     :|    |:
Db        560 GQDGRPGPPGPPGARGQAGVMGFPGPKGAAGEPGKAG-----ERGVPGPPGAVGPAGKDG 614

Qy       1294 GRSSSGPPHETAAPKRTYDMMEGRVGRAISSASIEGLMGRAIPP-ERHSPHHLKEQHHIR 1352
                : |||   | |         | | |  |   :|| | | || |   |
Db        615 EAGAQGPP-GPAGPAGE----RGEQGPA-GSPGFQGLPGPAGPPGEAGKPGE-------- 660

Qy       1353 GSITQGIPRSYVEAQEDYLRREAKLLKREGTPPPPPPSRDLTEAYKTQALGPLKLK--PA 1410
                 ||:|          | |       |   || |:            ||      |
Db        661 ----QGVPGDLGAPGPSGARGERGFPGERGVQGPPGPA-----------GPRGANGAPG 704

Qy       1411 HEGLVATVKEAGRSIHEIPREELRHTPELPLAPRPLKEGSITQGTP----LKYDTGASTT 1466
                ::|     :||                | ||     ||  || |    :  : ||:
Db        705 NDGAKG---DAG---------------APGAP-----GS--QGAPGLQGMPGERGAAGL 738

Qy       1467 GSKKHD-----VRSLIGSPGR------TFP--PVHPLDVMADARALERACYEESLKSRPG 1513
                 | |       :   ||||:       | |  |  |  |              |   |
Db        739 PGPKGDRGDAGPKGADGSPGKDGVRGLTGPIGPPGPAGAPGD------------KGESG 785

Qy       1514 TASSSGGSIARGAPVIVPELGKPRQSPLTYEDHGAPFAGHLPRGSPVTMREPTPRLQEGS 1573
                :  :| : |||||    | | |   |  :   | | |      | |    ||     :|
Db        786 PSGPAGPTGARGAPGDRGEPGPP--GPAGFA--GPPGAD----GQPGAKGEPGDAGAKGD 837

Qy       1574 LSSSKASQDRKLTSTPREIAKSPHSTVPEHHPHPI----SPYEHLLRGVSGVDLYRSHIP 1629
                                 | |   |  | ||    :|     || :|
Db        838 ------------------AGPPGPAGPAGPPGPIGNVGAPGAKGARGSAGP-------- 870

Qy       1630 LAFDPTSIPRGIPLDAAAAYYLPRHLAPNPTYPHLYPPYLIRGYPDTAALENRQTIINDY 1689
                        |       ||    |   | |:  :  ||    | | |  |
Db        871 --------PGATGFPGAAGRVGP----PGPS-GNAGPP----GPPGPAGKEG-------- 905

Qy       1690 ITSQQMHHNTATAMAQRADMLRGLSPRESSLALNYAAGPRGIIDLSQVPHLPVLVPPTPG 1749
                                 |  ||  :       || |          | | | ||
Db        906 ---------------------GKGPRGET-------GPAG---------RPGEVGP-PG 926

Qy       1750 TPATAMDRLAYLPTAPQPFSSRHSSSPLS------------------PGGPTHLTKPTT 1790
                 |  | :: :  | |  |  :  :  |                  || |    :|
Db        927 PPGPAGEKGS--PGADGPAGAPGTPGPQGIAGQRGVVGLPGQRGERGFPGLPGPSGEPGK 984

Qy       1791 TSSSERERDRDRERDRDREREKSILTSTTTVEHAPIWRPGTEQSSGSSGSSG-----GGG 1845
                    |    :|                    |: ||     | ||  |     |
Db        985 QGPSGASGERG--------------------PPGPMGPPGLAGPPGESGREGAPAAEGSP 1024

Qy       1846 GSSSRPASHSHAHQHSPISPRTQDALQQRPSVLHNTGMKGIITAVEPSKPTVLRSTSTSS 1905
                |    | :     :  | |           |  :   | |      |: |       :
Db       1025 GRDGSPGAKGDRGETGPAGPPGAPGAPGAPGPVGPAGKSGDRGETGPAGP--------AG 1076

Qy       1906 PVRPAATFPPATHCPLGGTLDGVYPTLMEPVLLPKEAPRVARPERPRADTGH-------- 1957
                || |     |                          | |: || | |
Db       1077 PVGPVGARGP-----------------------------AGPQGPRGDKGETGEQGDRG 1106

Qy       1958 -------AFLAKPPARSGLEPASSPSKGSEPRPLVPPVSGHATIARTPAKNLAPHHASPD 2010
                      : |  ||  |      ||  | |
Db       1107 IKGHRGFSGLQGPPGPPGSPGEQGPSGASGP--------------------------AG 1139

Qy       2011 PPAPPASASDPHREKTQSKPFSIQELELRSLGYHGSSYSPEGVEPVSPVSSPSLTHDKGL 2070
                |  || ||  | ::     |  |        |  |         || |  |
Db       1140 PRGPPGSAGAPGKDGLNGLPGPI--------GPPGPRGRTGDAGPVGPPGPPG------- 1184

Qy       2071 PKHLEELDKSHLEGELRPKQPGPVKLGGEAAHLPHLRPLPESQPSSSPLLQTAPGVKGHQ 2130
                                 |  |||   | : :||        ||          | | |
Db       1185 ----------------PPGPPGPPSAGFDFSFLP--------QP---------PQEKAHD 1211

Qy       2131 ----------------------RVVTLAQHISEVITQDYTRHHPQQLSAPLPAPLYSFPG 2168
                                    : :|:| |  : : : :| :| :
Db       1212 GGRYYRADDANVVRDRDLEVDTTLKSLSQQIENIRSPEGSRKNPAR-------------- 1257

Qy       2169 ASCPVLDLRRPPSDLYLPPPDHGAPARGSPHSEGGKRSPEPNKTSVLGG---------GE 2219
                :|   ||:   ||                    : |:   :||:   |          ||
Db       1258 -TC--RDLKMCHSDW----------------KSGEYWIDPNQGCNLDAIKVFCNMETGE 1297

Qy       2220 DGIEPVSPPEGM------TEPGHSRSAVYPLLYRDGEQTEPSRMGSKSPGNTSQPPAFFS 2273
                : | |          |   |   :    || | |          | | |
Db       1298 TCVYPTQPSVAQKNWYISKNPKDKRHVWFGESMTDGFQFEYG-------GQGSDPADVAI 1350

Qy       2274 KLTESNSAMVKSKKQEINKKLNTHNRNEPEYNISQPG 2310
                :||      ::    | :: :  | :|   |   | |
Db       1351 QLT-----FLRLMSTEASQNITYHCKNSVAYMDQQTG 1382

RESULT 67
US-10-177-293-65
; Sequence 65, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
;  APPLICANT: Lillie, James
;  APPLICANT:  Glatt, Karen
;  APPLICANT:  Zhao, Xumei
```

```
;  APPLICANT:  Gannavarpu, Manjula
;  APPLICANT:  Kamatkar, Shubhangi
;  APPLICANT:  Mertens, Maureen
;  APPLICANT:  Myer, Vic
;  APPLICANT:  Wang, Youzhen
;  APPLICANT:  Xu, Yongyao
;  APPLICANT:  Hoersch, Sebastian
;  APPLICANT:  Monahan, John
;  APPLICANT:  Meyers, Rachel E.
;  APPLICANT:  Bast Jr., Robert C.
;  APPLICANT:  Hortobagyi, Gabriel N.
;  APPLICANT:  Pusztai, Lajos
;  APPLICANT:  Meric, Funda
;  APPLICANT:  Sahin, Aysegul
;  APPLICANT:  Mills, Gordon B.
;  TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
;  TITLE OF INVENTION:  PREVENTION, AND THERAPY OF BREAST CANCER
;  FILE REFERENCE: MRI-038
;  CURRENT APPLICATION NUMBER: US/10/177,293
;  CURRENT FILING DATE:  2002-06-21
;  PRIOR APPLICATION NUMBER: US 60/299,887
;  PRIOR FILING DATE: 2001-06-21
;  PRIOR APPLICATION NUMBER: US 60/301,572
;  PRIOR FILING DATE: 2001-06-27
;  PRIOR APPLICATION NUMBER: US 60/306,501
;  PRIOR FILING DATE: 2001-07-18
;  PRIOR APPLICATION NUMBER: US 60/325,002
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US 60/362,585
;  PRIOR FILING DATE: 2002-03-05
;  PRIOR APPLICATION NUMBER: US 60/xxx,xxx
;  PRIOR FILING DATE: 2002-05-14
;  NUMBER OF SEQ ID NOS: 506
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 65
;   LENGTH: 1464
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-177-293-65

  Query Match             3.0%;  Score 392.5;  DB 14;  Length 1464;
  Best Local Similarity   21.3%;  Pred. No. 1.1e-09;
  Matches  378;  Conservative  102;  Mismatches  638;  Indels  659;  Gaps   84;

Qy        722 EALHASGNEVPRGECSGPATVNNSSDTESIPSPHTEAAK-DTGQNGPKPPA-TLGADGPP 779
              |  :  | ||| |||  |   : |      :    |  | |||  ||: ||   | || |
Db         77 ETKNCPGAEVPEGECC-PVCPDGSESPTDQETTGVEGPKGDTGPRGPRGPAGPPGRDGIP 135

Qy        780 -----PGPPTPPRRTSRAPIEPTPASEATGAPTPP----------------------P 810
                  |||| ||                    | | ||                    |
Db        136 GQPGLPGPPGPP----------------GPPGPPGLGGNFAPQLSYGYDEKSTGGISVP 178

Qy        811 APPSPSAP-----PPVVPKEEKEEETAAAPPVEEGE--------EQKPPAAEELAVDTGK 857
               |  || |      ||  |  :  :     || | ||         : ||       | |:
Db        179 GPMGPSGPRGLPGPPGAPGPQGFQ----GPPGEPGEPGASGPMGPRGPPGPPGKNGDDGE 234

Qy        858 AEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGSGRATTAKSSGAPQDSDS 917
              | :| :       | ||   : |    ||   |   |   |      ||   |
Db        235 AGKPGR-----PGERGPPGPQGARGLPGTA--GLPGMKGHRGFSGLDGAKGDAGP----- 282

Qy        918 SATCSADEVDEAEGGDKNRLLSP----RPSLLTPTGDPRANASPQKPLDLKQLKQRAAAI 973
                          | |    ||      |  : | | |      |  |          | |
Db        283 ------------AGPKGEPGSPGENGAPGQMGPRGLPGERGRPGAP-------GPAGAR 322

Qy        974 PPIQVTKVHEPPREDAAPTKPA-PPAPPPPQNLQPESDAPQQPGSS--PRGKSRSPAPP- 1029
                    |    ||     || || ||  |      : |:  || |  |  |:|    | ||
Db        323 GNDGATGAAGPP----GPTGPAGPPGFPGAVGAKGEA-GPQGPRGSEGPQGVRGEPGPPG 377

Qy       1030 ------------ADKEAFAAEAQKLPGDPPCWTSGLP-FPVPPREVIKASPHAPDPSAFS 1076
                         || :  |  |   ||      :| | ||          |  |     :
Db        378 PAGAAGPAGNPGADGQPGAKGANGAPG-----IAGAPGFPGARGPSGPQGPGGPPGPKGN 432

Qy       1077 YAPPGHPLPLGLHDTARPVLPRPPTISNPPPLISSAKHPSVLERQIGAISQGMSVQLHVP 1136
                                                                          |
Db        433 SGEPGAPGSKG--DTGAKGEPGPVGVQGPP-------GPAGEEGKRG------------ 470

Qy       1137 YSEHAKAPVGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPESLGVPTAQEASVLR 1196
                  |:   ||   |||                      | |: | | | | | |  : :
Db        471 ----ARGEPGPT--GLP------------------GPPGERGGPGSRGFPGADGVAGPK 505

Qy       1197 GTA--LGSVPGGSITKGIP-STRVPSDSAITYRGSITHGTPADVLYKGTITRIIGEDSPS 1253
                | |   || || :  || |      | :: :     :| |:|       |       |   |:
Db        506 GPAGERGS-PGPAGPKGSPGEAGRPGEAGLPGAKGLT-GSPGSPGPDGK----TGPPGPA 559

Qy       1254 RLD----------RGREDSL----PKGHVIYEGKKGHVLSYEGGM-----SVTQCSKED 1293
                |           ||:   :    |||      || |      | |:     :|    |:
Db        560 GQDGRPGPPGPPGARGQAGVMGFPGPKGAAGEPGKAG-----ERGVPGPPGAVGPAGKDG 614

Qy       1294 GRSSSGPPHETAAPKRTYDMMEGRVGRAISSASIEGLMGRAIPP-ERHSPHHLKEQHHIR 1352
                : |||   | |          | | |  |   :|| | | || |   |
Db        615 EAGAQGPP-GPAGPAGE----RGEQGPA-GSPGFQGLPGPAGPPGEAGKPGE-------- 660

Qy       1353 GSITQGIPRSYVEAQEDYLRREAKLLKREGTPPPPPPSRDLTEAYKTQALGPLKLK--PA 1410
                ||:|             | |        |   || |:            ||      |
Db        661 ----QGVPGDLGAPGPSGARGERGFPGERGVQGPPGPA-----------GPRGANGAPG 704

Qy       1411 HEGLVATVKEAGRSIHEIPREELRHTPELPLAPRPLKEGSITQGTP----LKYDTGASTT 1466
                ::|     :||                    | ||     ||  || |    :  : ||:
Db        705 NDGAKG---DAG---------------APGAP-----GS--QGAPGLQGMPGERGAAGL 738

Qy       1467 GSKKHD-----VRSLIGSPGR------TFP--PVHPLDVMADARALERACYEESLKSRPG 1513
                | |        :  ||||:      | |  |  |    |                |  |
Db        739 PGPKGDRGDAGPKGADGSPGKDGVRGLTGPIGPPGPAGAPGD------------KGESG 785

Qy       1514 TASSSGGSIARGAPVIVPELGKPRQSPLTYEDHGAPFAGHLPRGSPVTMREPTPRLQEGS 1573
                :  :| : |||||     | | |   |  :   | | |      | |    ||     :|
Db        786 PSGPAGPTGARGAPGDRGEPGPP--GPAGFA--GPPGAD----GQPGAKGEPGDAGAKGD 837

Qy       1574 LSSSKASQDRKLTSTPREIAKSPHSTVPEHHPHPI----SPYEHLLRGVSGVDLYRSHIP 1629
                                  |   |    |  ||    :|     || :|
Db        838 ------------------AGPPGPAGPAGPPGPIGNVGAPGAKGARGSAGP-------- 870

Qy       1630 LAFDPTSIPRGIPLDAAAAYYLPRHLAPNPTYPHLYPPYLIRGYPDTAALENRQTIINDY 1689
                         |       ||     |  | |: :  ||    | |  | |
Db        871 --------PGATGFPGAAGRVGP----PGPS-GNAGPP----GPPGPAGKEG-------- 905

Qy       1690 ITSQQMHHNTATAMAQRADMLRGLSPRESSLALNYAAGPRGIIDLSQVPHLPVLVPPTPG 1749
                                     |  ||  :      || |          | | | ||
Db        906 ----------------------GKGPRGET-------GPAG---------RPGEVGP-PG 926

Qy       1750 TPATAMDRLAYLPTAPQPFSSRHSSSPLS------------------PGGPTHLTKPTT 1790
                |  | :: :  | |  |  :  :  |                   || |     :|
Db        927 PPGPAGEKGS--PGADGPAGAPGTPGPQGIAGQRGVVGLPGQRGERGFPGLPGPSGEPGK 984

Qy       1791 TSSSERERDRDRERDRDRERERKSILTSTTTVEHAPIWRPGTEQSSGSSGSSG-----GGG 1845
                |    :|                      |:  ||    | ||  |     |
Db        985 QGPSGASGERG------------------PPGPMGPPGLAGPPGESGREGAPAAEGSP 1024

Qy       1846 GSSSRPASHSHAHQHSPISPRTQDALQQRPSVLHNTGMKGIITAVEPSKPTVLRSTSTSS 1905
                |    | :     :  | |          |  :   | |       |: |       :
Db       1025 GRDGSPGAKGDRGETGPAGPPGAPGAPGAPGPVGPAGKSGDRGETGPAGP--------AG 1076

Qy       1906 PVRPAATFPPATHCPLGGTLDGVYPTLMEPVLLPKEAPRVARPERPRADTGH-------- 1957
                || |   |                          | |: || | |
Db       1077 PVGPVGARGP----------------------------AGPQGPRGDKGETGEQGDRG 1106

Qy       1958 -------AFLAKPPARSGLEPASSPSKGSEPRPLVPPVSGHATIARTPAKNLAPHHASPD 2010
                    : |  ||   |      ||  | |
Db       1107 IKGHRGFSGLQGPPGPPGSPGEQGPSGASGP-------------------------AG 1139

Qy       2011 PPAPPASASDPHREKTQSKPFSIQELELRSLGYHGSSYSPEGVEPVSPVSSPSLTHDKGL 2070
                |  || ||  | ::     |  |       |  |          || |   |
Db       1140 PRGPPGSAGAPGKDGLNGLPGPI--------GPPGPRGRTGDAGPVGPPGPPG------- 1184

Qy       2071 PKHLEELDKSHLEGELRPKQPGPVKLGGEAAHLPHLRPLPESQPSSSPLLQTAPGVKGHQ 2130
                                  |  |||   | : : ||        ||         |  | |
```

```
Db       1185 ----------------PPGPPGPPSAGFDFSFLP--------QP---------PQEKAHD 1211

Qy       2131 ----------------------RVVTLAQHISEVITQDYTRHHPQQLSAPLPAPLYSFPG 2168
                                    : :|:| |  : : : :| :| :
Db       1212 GGRYYRADDANVVRDRDLEVDTTLKSLSQQIENIRSPEGSRKNPAR-------------- 1257

Qy       2169 ASCPVLDLRRPPSDLYLPPPDHGAPARGSPHSEGGKRSPEPNKTSVLGG---------GE 2219
              :|   ||:   ||                  : |:   :||:   |            ||
Db       1258 -TC--RDLKMCHSDW----------------KSGEYWIDPNQGCNLDAIKVFCNMETGE 1297

Qy       2220 DGIEPVSPPEGM------TEPGHSRSAVYPLLYRDGEQTEPSRMGSKSPGNTSQPPAFFS 2273
                : |  |              |   |   :     || | |           |  | |
Db       1298 TCVYPTQPSVAQKNWYISKNPKDKRHVWFGESMTDGFQFEYG-------GQGSDPADVAI 1350

Qy       2274 KLTESNSAMVKSKKQEINKKLNTHNRNEPEYNISQPG 2310
              :||      ::    | :: :  | :|   |   | |
Db       1351 QLT-----FLRLMSTEASQNITYHCKNSVAYMDQQTG 1382


RESULT 68
US-10-301-822-28
; Sequence 28, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
;  APPLICANT: Millennium Pharmaceuticals, Inc.
;  APPLICANT:  Berger, Allison
;  APPLICANT:  Guillemette, Tracy L.
;  APPLICANT:  Kamatkar, Shubhangi
;  APPLICANT:  Schlegel, Robert
;  APPLICANT:  Monahan, John E.
;  APPLICANT:  Thibodeau, Stephen N.
;  APPLICANT:  Burgart, Lawrence J.
;  TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
;  TITLE OF INVENTION:  METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
;  TITLE OF INVENTION:  THERAPY OF COLON CANCER
;  FILE REFERENCE: MPM01-029P2RNM
;  CURRENT APPLICATION NUMBER: US/10/301,822
;  CURRENT FILING DATE:  2002-11-21
;  PRIOR APPLICATION NUMBER: US 60/339,971
;  PRIOR FILING DATE: 2001-12-10
;  PRIOR APPLICATION NUMBER: US 60/361,978
;  PRIOR FILING DATE: 2002-03-05
;  PRIOR APPLICATION NUMBER: US 60/381,988
;  PRIOR FILING DATE: 2002-05-20
;  NUMBER OF SEQ ID NOS: 228
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 28
;   LENGTH: 1464
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-10-301-822-28

  Query Match             3.0%;  Score 392.5;  DB 14;  Length 1464;
  Best Local Similarity   21.3%;  Pred. No. 1.1e-09;
  Matches  378;  Conservative  102;  Mismatches  638;  Indels  659;  Gaps   84;

Qy        722 EALHASGNEVPRGECSGPATVNNSSDTESIPSPHTEAAK-DTGQNGPKPPA-TLGADGPP 779
              |  :  | ||| |||  |   : |        :   |  | |||  ||: ||   | || |
Db         77 ETKNCPGAEVPEGECC-PVCPDGSESPTDQETTGVEGPKGDTGPRGPRGPAGPPGRDGIP 135

Qy        780 -----PGPPTPPRRTSRAPIEPTPASEATGAPTPP-----------------------P 810
                  |||| ||                  | | ||                       |
Db        136 GQPGLPGPPGPP----------------GPPGPPGLGGNFAPQLSYGYDEKSTGGISVP 178

Qy        811 APPSPSAP-----PPVVPKEEKEEETAAAPPVEEGE--------EQKPPAAEELAVDTGK 857
              |  || |     ||  |  :  :     || | ||         : ||       | |:
Db        179 GPMGPSGPRGLPGPPGAPGPQGFQ----GPPGEPGEPGASGPMGPRGPPGPPGKNGDDGE 234

Qy        858 AEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGSGRATTAKSSGAPQDSDS 917
              | :| :     | ||   : |     ||   |   |   |     ||    |
Db        235 AGKPGR-----PGERGPPGPQGARGLPGTA--GLPGMKGHRGFSGLDGAKGDAGP----- 282

Qy        918 SATCSADEVDEAEGGDKNRLLSP----RPSLLTPTGDPRANASPQKPLDLKQLKQRAAAI 973
```

```
                  | |    ||     | : | | |     | |       | |
Db        283 ------------AGPKGEPGSPGENGAPGQMGPRGLPGERGRPGAP-------GPAGAR 322

Qy        974 PPIQVTKVHEPPREDAAPTKPA-PPAPPPPQNLQPESDAPQQPGSS--PRGKSRSPAPP- 1029
                  |    ||     || || ||  |     : |:  || |  |  |:|    | ||
Db        323 GNDGATGAAGPP----GPTGPAGPPGFPGAVGAKGEA-GPQGPRGSEGPQGVRGEPGPPG 377

Qy       1030 ------------ADKEAFAAEAQKLPGDPPCWTSGLP-FPVPPREVIKASPHAPDPSAFS 1076
                         || :  |  |   ||      :| | ||           | |     :
Db        378 PAGAAGPAGNPGADGQPGAKGANGAPG-----IAGAPGFPGARGPSGPQGPGGPPGPKGN 432

Qy       1077 YAPPGHPLPLGLHDTARPVLPRPPTISNPPPLISSAKHPSVLERQIGAISQGMSVQLHVP 1136
                 || |  |  ||     | |  :  ||          |:  | : |
Db        433 SGEPGAPGSKG--DTGAKGEPGPVGVQGPP-------GPAGEEGKRG------------- 470

Qy       1137 YSEHAKAPVGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPESLGVPTAQEASVLR 1196
                 |:   ||   |||                    | |: | | | | | |   :  :
Db        471 ----ARGEPGPT--GLP------------------GPPGERGGPGSRGFPGADGVAGPK 505

Qy       1197 GTA--LGSVPGGSITKGIP-STRVPSDSAITYRGSITHGTPADVLYKGTITRIIGEDSPS 1253
                 | |   || ||  :  || |      | :: :     :| |:|      |      |  |:
Db        506 GPAGERGS-PGPAGPKGSPGEAGRPGEAGLPGAKGLT-GSPGSPGPDGK----TGPPGPA 559

Qy       1254 RLD----------RGREDSL----PKGHVIYEGKKGHVLSYEGGM-----SVTQCSKED 1293
                 |          ||:   :    |||     || |      | |:     :|    |:
Db        560 GQDGRPGPPGPPGARGQAGVMGFPGPKGAAGEPGKAG-----ERGVPGPPGAVGPAGKDG 614

Qy       1294 GRSSSGPPHETAAPKRTYDMMEGRVGRAISSASIEGLMGRAIPP-ERHSPHHLKEQHHIR 1352
                 : |||   | |          | | |  |   :|| | | || |   |
Db        615 EAGAQGPP-GPAGPAGE----RGEQGPA-GSPGFQGLPGPAGPPGEAGKPGE-------- 660

Qy       1353 GSITQGIPRSYVEAQEDYLRREAKLLKREGTPPPPPPSRDLTEAYKTQALGPLKLK--PA 1410
                 ||:|           | |       |   || |:             ||      |
Db        661 ----QGVPGDLGAPGPSGARGERGFPGERGVQGPPGPA-----------GPRGANGAPG 704

Qy       1411 HEGLVATVKEAGRSIHEIPREELRHTPELPLAPRPLKEGSITQGTP----LKYDTGASTT 1466
                 ::|     :||                 | ||      || || |    :  : ||:
Db        705 NDGAKG---DAG---------------APGAP-----GS--QGAPGLQGMPGERGAAGL 738

Qy       1467 GSKKHD-----VRSLIGSPGR------TFP--PVHPLDVMADARALERACYEESLKSRPG 1513
                 | |      :   ||||:       | |  |  |  |          |      | |
Db        739 PGPKGDRGDAGPKGADGSPGKDGVRGLTGPIGPPGPAGAPGD------------KGESG 785

Qy       1514 TASSSGGSIARGAPVIVPELGKPRQSPLTYEDHGAPFAGHLPRGSPVTMREPTPRLQEGS 1573
                 :  :| : |||||     | | |   |  :   | | |     | |    ||     :|
Db        786 PSGPAGPTGARGAPGDRGEPGPP--GPAGFA--GPPGAD----GQPGAKGEPGDAGAKGD 837

Qy       1574 LSSSKASQDRKLTSTPREIAKSPHSTVPEHHPHPI----SPYEHLLRGVSGVDLYRSHIP 1629
                                  | |   |  | ||    :|     || :|
Db        838 ------------------AGPPGPAGPAGPPGPIGNVGAPGAKGARGSAGP-------- 870

Qy       1630 LAFDPTSIPRGIPLDAAAAYYLPRHLAPNPTYPHLYPPYLIRGYPDTAALENRQTIINDY 1689
                    |       ||    |  |: :  ||    | |  |  |
Db        871 --------PGATGFPGAAGRVGP----PGPS-GNAGPP----GPPGPAGKEG-------- 905

Qy       1690 ITSQQMHHNTATAMAQRADMLRGLSPRESSLALNYAAGPRGIIDLSQVPHLPVLVPPTPG 1749
                                  |  ||  :      || |         | | | ||
Db        906 ----------------------GKGPRGET-------GPAG---------RPGEVGP-PG 926

Qy       1750 TPATAMDRLAYLPTAPQPFSSRHSSSPLS-------------------PGGPTHLTKPTT 1790
                 |  | :: :  | |  |  :  :  |                   | |  |    :|
Db        927 PPGPAGEKGS--PGADGPAGAPGTPGPQGIAGQRGVVGLPGQRGERGFPGLPGPSGEPGK 984

Qy       1791 TSSSERERDRDRERDRDREREKSILTSTTTVEHAPIWRPGTEQSSGSSGSSG-----GGG 1845
                    | :|                    |:  ||    | ||  |     |
Db        985 QGPSGASGERG-------------------PPGPMGPPGLAGPPGESGREGAPAAEGSP 1024

Qy       1846 GSSSRPASHSHAHQHSPISPRTQDALQQRPSVLHNTGMKGIITAVEPSKPTVLRSTSTSS 1905
                 |    | :     :  | |           | :   | |       |: |      :
Db       1025 GRDGSPGAKGDRGETGPAGPPGAPGAPGAPGPVGPAGKSGDRGETGPAGP--------AG 1076

Qy       1906 PVRPAATFPPATHCPLGGTLDGVYPTLMEPVLLPKEAPRVARPERPRADTGH-------- 1957
                 || |    |                          | |: || | |
```

```
Db       1077 PVGPVGARGP----------------------------AGPQGPRGDKGETGEQGDRG 1106

Qy       1958 -------AFLAKPPARSGLEPASSPSKGSEPRPLVPPVSGHATIARTPAKNLAPHHASPD 2010

Db       1107 IKGHRGFSGLQGPPGPPGSPGEQGPSGASGP--------------------------AG 1139

Qy       2011 PPAPPASASDPHREKTQSKPFSIQELELRSLGYHGSSYSPEGVEPVSPVSSPSLTHDKGL 2070

Db       1140 PRGPPGSAGAPGKDGLNGLPGPI--------GPPGPRGRTGDAGPVGPPGPPG------- 1184

Qy       2071 PKHLEELDKSHLEGELRPKQPGPVKLGGEAAHLPHLRPLPESQPSSSPLLQTAPGVKGHQ 2130

Db       1185 ----------------PPGPPGPPSAGFDFSFLP--------QP---------PQEKAHD 1211

Qy       2131 ----------------------RVVTLAQHISEVITQDYTRHHPQQLSAPLPAPLYSFPG 2168

Db       1212 GGRYYRADDANVVRDRDLEVDTTLKSLSQQIENIRSPEGSRKNPAR-------------- 1257

Qy       2169 ASCPVLDLRRPPSDLYLPPPDHGAPARGSPHSEGGKRSPEPNKTSVLGG---------GE 2219

Db       1258 -TC--RDLKMCHSDW----------------KSGEYWIDPNQGCNLDAIKVFCNMETGE 1297

Qy       2220 DGIEPVSPPEGM------TEPGHSRSAVYPLLYRDGEQTEPSRMGSKSPGNTSQPPAFFS 2273

Db       1298 TCVYPTQPSVAQKNWYISKNPKDKRHVWFGESMTDGFQFEYG-------GQGSDPADVAI 1350

Qy       2274 KLTESNSAMVKSKKQEINKKLNTHNRNEPEYNISQPG 2310

Db       1351 QLT-----FLRLMSTEASQNITYHCKNSVAYMDQQTG 1382
```

RESULT 69
US-10-029-386-33090
; Sequence 33090, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33090
; LENGTH: 1633
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC027307.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.85
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.84
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.8
; OTHER INFORMATION: SWISSPROT HIT: Q61315, EVALUE 1.00e-27
US-10-029-386-33090

Query Match 3.0%; Score 391.5; DB 14; Length 1633;
Best Local Similarity 20.5%; Pred. No. 1.4e-09;
Matches 383; Conservative 170; Mismatches 647; Indels 665; Gaps 88;

```
Qy        680 RRKKKKAPAAASEEAAFPPVVEDEEMEASGVSGNEEEMVEEAEALHAS------------ 727

Db        153 RRGGKEAEKDTSGEAA---VAAKAKAKLALAVARIDQLVEDISALHTSSDDSFSLSSGDP 209

Qy        728 GNEVPR-------GECSGPATVNNSSDTESIP-----SPHTEAAKDTGQNG--------- 766

Db        210 GQEAPREGRAQSCSPCRGPEGGRREAGSRAHPLLRLKAAHASLSNDSLNSGSASDGYCPR 269

Qy        767 ----PKPPATLGA--DGPPPGPPTPPRRTSRAP---IEPTPASEATGA----------PT 807

Db        270 EHMLPCPLAALASRREDPRCGQPRPSRLDLDLPGCQAEP-PAREATSADARVRTIKLSPT 328

Qy        808 PPPAP---------PSPSAPPPVVPKEEKE-----------EETAAAPPVEEGEEQKPP 846

Db        329 YQHVPLLEGASRAGAEPLAGPGISPGARKQAWLPADHLSKVPEKLAAAPLSVASKALQKL 388

Qy        847 AAEELAVDTGKAEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGSGRATTA 906

Db        389 AAQEGPLSLSRCSS--LSSLSSAGRPGPSEGGDLDDSDSSLEGL-----EEAGPSEAELD 441

Qy        907 KSSGAPQDSDSSATCSADEVDEAEG-GDKNRLLSPRPSLLTPTGDPRANASPQKPLDLKQ 965

Db        442 STWRAPGATSLPVAIPAPRRNRGRGLGVED---------ATPSSSSE-NYVQETPLVLSR 491

Qy        966 LKQRAAAIPPIQVTKVHEPPREDAAPTKPAPPAPPPPQNLQPESDAPQQPGSS-PRGKSR 1024

Db        492 CS---------SVSSLGSFESPSIASSIPSEPCSGQGSGTISPSELPDSPGQTMPPSRSK 542

Qy       1025 ----SPAPPADKEAFAAEAQKLPGDPPCWTS---------------GLPFPVPPREVIKA 1065

Db        543 TPPLAPAPQGPPEATQFSLQ--------WESYVKRFLDIADCRERCRLPSELDAGSV-RF 593

Qy       1066 SPHAPDPSAFSYAPPGHPLPLGLHDTARPVLPRPPTISNPPPLISSAKHPSVLERQIGAI 1125

Db        594 TVEKPDEN-FSCASSLSALALHEHYVQQDVELR---------LLPSA----CPERGGGAG 639

Qy       1126 SQGMSVQLHVPYSEHAKAPVGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPESLG 1185

Db        640 GAG----LH--FAGHRRREEGPAPTG--------------------SRPRGAAD------ 667

Qy       1186 VPTAQEASVLRGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHGTPADVLYKGTITR 1245

Db        668 ----QELELLRECLGAAVP--ARLRKVASALVPGRRAL----------PVPVY------M 705

Qy       1246 IIGEDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPHETA 1305

Db        706 LVPAPAPAQEDDSCTDS----------AEGTPVNFSSAASLSD------ETLQGPPR--- 746

Qy       1306 APKRTYDMMEGRVGRAISSASIEGLMGRAIPPERHSPHHLKEQHHIRGSITQGIPRSYVE 1365

Db        747 ------DQPGGPAGR-------QRPTGRPTSARQAMGHRHK---------AGGAGRS-AE 783

Qy       1366 AQEDYLRREAKLLKREGTPPPPPSRDLTEAYKTQALGPLK----LKPAHEGL------- 1414

Db        784 QSRGAGKNRAGLELPLGRPPSAPADKDGSKPGRTRGDGALQSLCLTTPTEEAVYCFYGND 843

Qy       1415 ------VATVKEAGRSIHEIPREELRHTPE----------LPLAPRPLK--EGSITQGT 1455

Db        844 SDEEPPAAAPTPTHRRTSAIPRAFTRERPQGRKEAPAPSKAAPAAPPPARTQPSLIADET 903

Qy       1456 PLKYDTGASTTGSKKHDVRSLIGSPGRTFPP-VHP----------------LDVMADARA 1498

Db        904 PPCYSLSSSA---------SSLSEPEPSEPPAVHPRGREPAVTKDPGPGGGRDSSPSPRA 954

Qy       1499 LE---RACYEES--------------------LKSRPGTASSSGG--------------- 1520

Db        955 AEELLQRCISSALPRRRPPVSGLRRRKPRATRLDERPAEGSRERGEEAAGSDRASDLDSV 1014

Qy       1521 ---SIARGAPVIVP---ELGKPRQSPLTYEDHGAPFAGHLPRGSPVTMREP---TPRLQE 1571

Db       1015 EWRAIQEGANSIVTWLHQAAAATREASSESDSILSFVSGLSVGS--TLQPPKHRKGRQAE 1072

Qy       1572 GSLSSSKASQDR-----KLTSTPREIA--KSPHST------VPEHHPHPISPYEHLLRGV 1618

Db       1073 GEMGSARRPEKRGAASVKTSGSPRSPAGPEKPRGTQKTTPGVPA----------VLRGR 1121

Qy       1619 SGVDLYRSHIPLAFDPTSIPRGIPLDAAAAYYLPRHLAPNPTYPHLYPPYLIRGYPDTAA 1678

Db       1122 T----------VIYVPSPAPRAQPKGTPGPRATPRKVAP-------------------- 1150

Qy       1679 LENRQTIINDYITSQQMHHNTATAMAQRADMLRGLSP-RESSLALNYAAGPRGIIDLSQV 1737

Db       1151 ----------------------PCLAQPAAPAKVPSPGQQRSRSLHRPAKTSELATLSQ- 1187
```

```
Qy       1738 PHLPVLVPPTPGTPATAMDRLAYLPTAPQPFSSRHSSSPLS------------------- 1778
                                                                          
Db       1188 -------PPRSATPPARL--------AKTPSSSSSQTSPASQPLPRKRPPVTQAAGALPG 1232

Qy       1779 PG-GPTHLTKPTTTSSSERERDRDRER----DRDRREREKSILTSTTTVEHAPIWRPGTEQ 1833
                                                                          
Db       1233 PGASPVPKTPARTLLAKQHKTQRSPVRIPFMQRPARRGPPPL-ARAVPEPGPRGRAGTEA 1291

Qy       1834 SSGSSGSSGGGGGSSSRPASHSHAHQHSPISPRTQDALQQRPSVLHNTGMKGIITAVEPS 1893
                                                                          
Db       1292 GPGARGGRLG--------------------LVRVASALSSGSESSDRSGFRRQLTFIKES 1331

Qy       1894 KPTVLRSTSTSSPVRPAATFPPATHCPLGGTLDGVYPTLMEPVL---------------L 1938
                                                                          
Db       1332 -PGLRRRRSELSSAESAASAP-----------QGASPRRGRPALPAVFLCSSRCEELRAA 1379

Qy       1939 PKEAPRVARPERPRADTGHAFLAKPPARSGLEPASSPSKGSEPR-----------PLVPP 1987
                                                                          
Db       1380 PRQGPAPAR-QRP------------------PAARPSPGERPARRTTSESPSRLPVRAP 1419

Qy       1988 VSGHATIARTPAKNLAPHHA---SPD--PPAPPASASDPHREKTQSKPFSIQELELRSLG 2042
                                                                          
Db       1420 AARPETVKRYAS---LPHISVARRPDGAVPAAPASA-DAARRSSDGEP------------ 1463

Qy       2043 YHGSSYSPEGVEPVSPVSSPSLTHDKGLPKHLEELDKSHLEGELRPKQPGPVKLGGEAAH 2102
                                                                          
Db       1464 -----------RPLPRVAAPGTTW-----RRIRDEDVPHI---LRSTLP--------ATA 1496

Qy       2103 LPHLRPLPESQPSSSPLLQTAPGVKGHQRVVTLAQHISEVITQDYTRHHPQQLSAPLPAP 2162
                                                                          
Db       1497 LPLRGSTPEDAPAGPPPRKT-----------------SDAVVQ------TEEVAAP---- 1529

Qy       2163 LYSFPGASCPVLDLRRPPSDLYLPPPDHGAPARGSPHSEGGKRSPEPNKTSVLGGGEDGI 2222
                                                                          
Db       1530 --KTNSSTSPSLETREPP----------GAPAGG-------------QLSLLGSDVDGP 1563

Qy       2223 E----PVSPP-----EGMTEPGHSRSAVYPLLYRDGEQTEPSRMGSKSPGNTSQPPAFFS 2273
                                                                          
Db       1564 SLAKAPISAPFVHEGLGVAVGGFPAS-------RHG---SPSRSARVPPFNYVPSPMVVA 1613

Qy       2274 KLTES 2278
                 
Db       1614 ATTDS 1618
```

RESULT 70
US-10-052-648A-20
; Sequence 20, Application US/10052648A
; Publication No. US20040005558A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David
; APPLICANT: Burgess, Catherine
; APPLICANT: Casman, Stacie
; APPLICANT: Colman, Steven
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Gerlach, Valerie
; APPLICANT: Gunther, Erik
; APPLICANT: Kekuda, Ramesh
; APPLICANT: MacDougall, John R.
; APPLICANT: Mehraban, Fuad
; APPLICANT: Patturajan, Meera
; APPLICANT: Rothenberg, Mark
; APPLICANT: Shimkets, Richard
; APPLICANT: Smithson, Glennda
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Stone, David J.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Zerhusen, Bryan D.
; TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
; TITLE OF INVENTION: USING THE SAME
; FILE REFERENCE: 21402-250 (CURA-550)
; CURRENT APPLICATION NUMBER: US/10/052,648A
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: 60/262,454
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/272,920
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/284,549
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/303,229
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/262,892
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/263,605
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/269,098
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: 60/264,159
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/271,855
; PRIOR FILING DATE: 2001-02-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 2058
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-052-648A-20

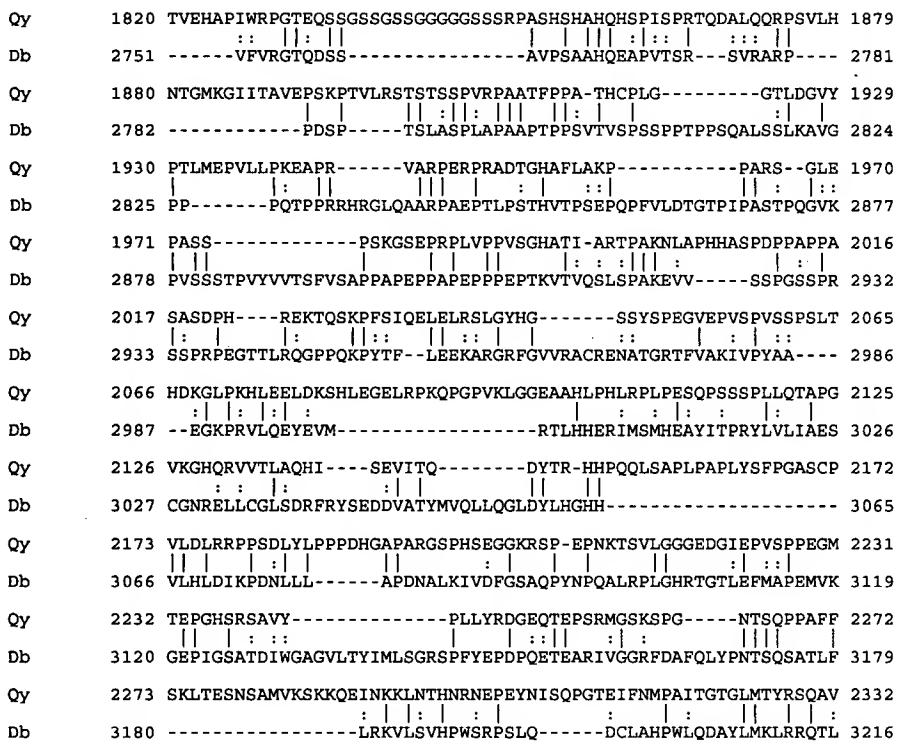
Query Match 2.9%; Score 388.5; DB 15; Length 2058;
Best Local Similarity 21.5%; Pred. No. 2.6e-09;
Matches 365; Conservative 196; Mismatches 604; Indels 533; Gaps 87;

```
Qy         50 ASHLSPGSIIQPQRRRPSLLSEFQPGNE---------RSQELHLRPESHSYLPELGKSEM 100
                                                                          
Db        443 AHTILPGS----KTREPSNAKQQKEKNKVKKKRKKETKGREVRFGKRSDKFCS--SESES 496

Qy        101 EFIES-KRPRLELLPDPLLRPSPLLATGQPAGSEDLTKDRSLTGKLEPVSPPS------- 152
                                                                          
Db        497 ESSESGEDDRDSLGSSGCLKGSPL-----------VLKDPSLFSSLSASSTSSHGSSAAQ 545

Qy        153 ---PPHTDPELELVPPRLSKEELIQNMDRV----DREITMVEQQISKLK--KKQQQLEEE 203
                                                                          
Db        546 KQNDQHTKHWKTISSPAWSEVSSLSDSTRTRLTSESDYSSEGSSVESLKPVRKRQEHRKR 605

Qy        204 AA----KPPEPEKPVSPPPIESKHRSLVQIIYDENRKKAEAAHRILEGLGPQVELPLY-- 257
                                                                          
Db        606 ASLSEKKSPFLSSAEGAVPKLDKEGKVVK------KHKTKHKHKNKEKISQELKLKSFTY 659

Qy        258 -----NQPSDTRQYHENIKINQAMRKKLILYFKRRNHARKQWK-QKFCQRYDQLMEALEK 311
                                                                          
Db        660 EYEDSKQKSDKAILLEN---DLSTENKLKVLKHDRDHFKKEEKLSKMKLEEKEWLFKDEK 716

Qy        312 KVERIENNPRRRAKE-SKVREYYEKQFPE----IRKQRELQE----RMQSRVGQRGSGLS 362
                                                                          
Db        717 SLKRIKDKLRLYKEERDKISKEKEKIFKEDKEKLKKEKVYREDSLSDRDSSFDFKGAKLI 776

Qy        363 M-------------SAARSEHEVSEIIDGLSEQENLEKQMRQLAVIPPMLYDADQQRIKF 409
                                                                          
Db        777 LETVKEDSKERRRDSRAREKHPARE-----KEKPDKRKRYKE--------KDKDKSEKSI 823

Qy        410 INMNGLMADPMKVYKDRQVMNMWSEQEKETFREKFMQHPKNFGLIASFLERKTVAECVLY 469
                                                                          
Db        824 LE---------KCQKDKEK----KEKHKDT-------HGKDKERKASVFEK--------- 854

Qy        470 YYLTKKN----ENYKSLVRRSYRRRGKSQQQQQQQQQQQQQQQQQQPMPRSSQEEKDEKEK 525
                                                                          
Db        855 -HKEKKDKESTEKYKD--RASV---DSTQDKKNKQEKAEKKHAAEDKAKSKHKEKSDKEH 908

Qy        526 EKE------AEKEEEKPEVEN----DKED--------LLKEK----TDDTSGEDNDEKEA 563
                                                                          
Db        909 SKERKSSRSADAEYRESEVSSDSFTDREDDKSACLPEKLKEKRHRHSSSSSKKSHDREE- 967
```

```
Qy       564 VASKGRKTANSQGRRKGRITRSMANEANSEEAITPQQSAELASMELNESSRWTEEEMETA 623
Db       968 -----KKEDYKEG-RKGQYEKDL-------EADAYGVSYNMKAIELFEKKDKNDEPLKEK 1014

Qy       624 KKGLLEHGRNW----------------------------------------SAIARMVGS 643
Db      1015 KK-REKHREKWRDEKERHRDRHADRPKPSKDPGKKDARPREKLLGDGDLMMTSFERMLSQ 1073

Qy       644 KTVSQCKNFYFNYKKRQNLDEILQQH----KLKMEKERNA--RRKK------KKAPA--- 688
Db      1074 KDL-EIEERHKRHKERMKQMEKLRHRSGDPKLK-EKAKPADDGRKKGLDIPAKKPPGLDP 1131

Qy       689 -----AASEEAAFPPVVEDEEMEASGVSGNE----EEMVEEAEALHASGNEVPRGECSGP 739
Db      1132 PFKDKKLKESTPIPPAAENKLHPASGADSKDWLAGPHM---KEVLPAS----PRPDQSRP 1184

Qy       740 ATVNNSSDTESIPS-------PHTEAAK------------DTGQNGPKPPATLGADGP-- 778
Db      1185 TGVPTPTSVLSCPSYEEVMHTPRTPSCSADDYADLVFDCADSQHSTPVPTAPTSACSPSF 1244

Qy       779 ------------------PPGP-----------------------------PTPPRRTS 790
Db      1245 FDRFSVASSGLSENASQAPARPLSTNLYRSVSVDIDKLFRQQSVPAASSYDSPMPPSMED 1304

Qy       791 RAPIEPTPA-----------SEATGAPTP-------PPA-----PPS-----------PS 816
Db      1305 RAPLPPVPAEKFACLSPGYYSPDYGLPSPKVDALHCPPAAVVTVTPSPEGVFSSLQAKPS 1364

Qy       817 APPPVVPKEEKEEETAAAPPVEEGEEQKPPAAEELAVDTGKAEEPVKSECTEEAEEGPAK 876
Db      1365 PSPPSLDTSEDQQATAAIIPPE------PSYLEPL------DEGPFSAVITEEPVEWAHP 1412

Qy       877 GKDAEAAEATAEGALKAEKKEGGSGRATTAKSSG------APQDSDSSAT--CSADEVDE 928
Db      1413 SEQALASSLI-----------GGTSENPVSWPVGSDLLLKSPQRFPESPKRFCPADPLHS 1461

Qy       929 AEGGDKNRLLSPRPSLLTPTGDPRANASPQKPLDLKQLKQRAAAIPPIQVTKVHEPPRED 988
Db      1462 AAPGPFSASEAPYPA---PPASPAPYALP--VAELEDVKDVPAAI-----------STSE 1505

Qy       989 AAPTKPAPPA------------PPPPQNLQPESDAPQQPG--SSPRGKSR----SPAPPA 1030
Db      1506 AAPY--APPSGLESFFSNCKSLPEAPLDVAPEALGPLENSFLDGSRGLSHLGQVEPVPWA 1563

Qy      1031 DKEAFAAEAQKLPGDPPCWTSGLPFPVPPREVIKASPHAPDPSAFSYAPPGHPLPLGLHD 1090
Db      1564 D--AFAGPEDDLDLGP------FSLPELPLQTKDAADGEAEPVEESLAPP-EEMPPG--- 1611

Qy      1091 TARPVLPRPPTISNPPPLISSAKHPSVLERQIGAISQGMSVQLHVPYSEHAKAPVGPVTM 1150
Db      1612 APRELEPEP---SGEPKL------DVALEAAVEAETVPEERARGDPDSSVEPAPVPPEQL 1662

Qy      1151 GLPLPMDPKKLAPFSGVKQEQLSPR--GQAGPPESLGVPTAQEASVLRGTALGSVPGGSI 1208
Db      1663 G---SGDPSLCAP------DGPAPNTVAQAQAADGAGPEDDTEAS--RAAAPAEGPPGQ- 1710

Qy      1209 TKGIPSTRVPSDSAITYRGSITHGTPADVLYKGTITRIIGEDSPSRLDRGREDSLPKGHV 1268
Db      1711 ----PEAAEPKPTAEAPK------APREI-------------PQRMTRNRAQMLAN--- 1743

Qy      1269 IYEGKKGHVLS-YEGGMSVTQCSKEDGRSSSGPPHETAAPKR-------------TYDMM 1314
Db      1744 --QSKQGPPPSEKECAPTPAPVTRAKARGSEDDDAQAQHPRKRRFQRSTQQLQLNTSTQQ 1801

Qy      1315 EGRVGRAISSASIEGLMGRAIPP---ERHSPH--------HLKEQHHIRGSITQGIPRSY 1363
Db      1802 TREVIQQTLAAIVDAIKLDAIEPYHSDRANPYFEYLQIRKKIEEKRKILCCITPQAPQCY 1861

Qy      1364 VE---AQEDYLRREAKLLKREGTP---PPPPPSRDLTEAYKTQ--ALGPLKLKPAHEGLV 1415
Db      1862 AEYVTYTGSYL-LDGKPLSKLHIPVIAPPPSLAEPLKELFRQQEAVRGKLRLQ------- 1913

Qy      1416 ATVKEAGRSIHEIPREEL 1433
Db      1914 ----------HSIEREKL 1921
```

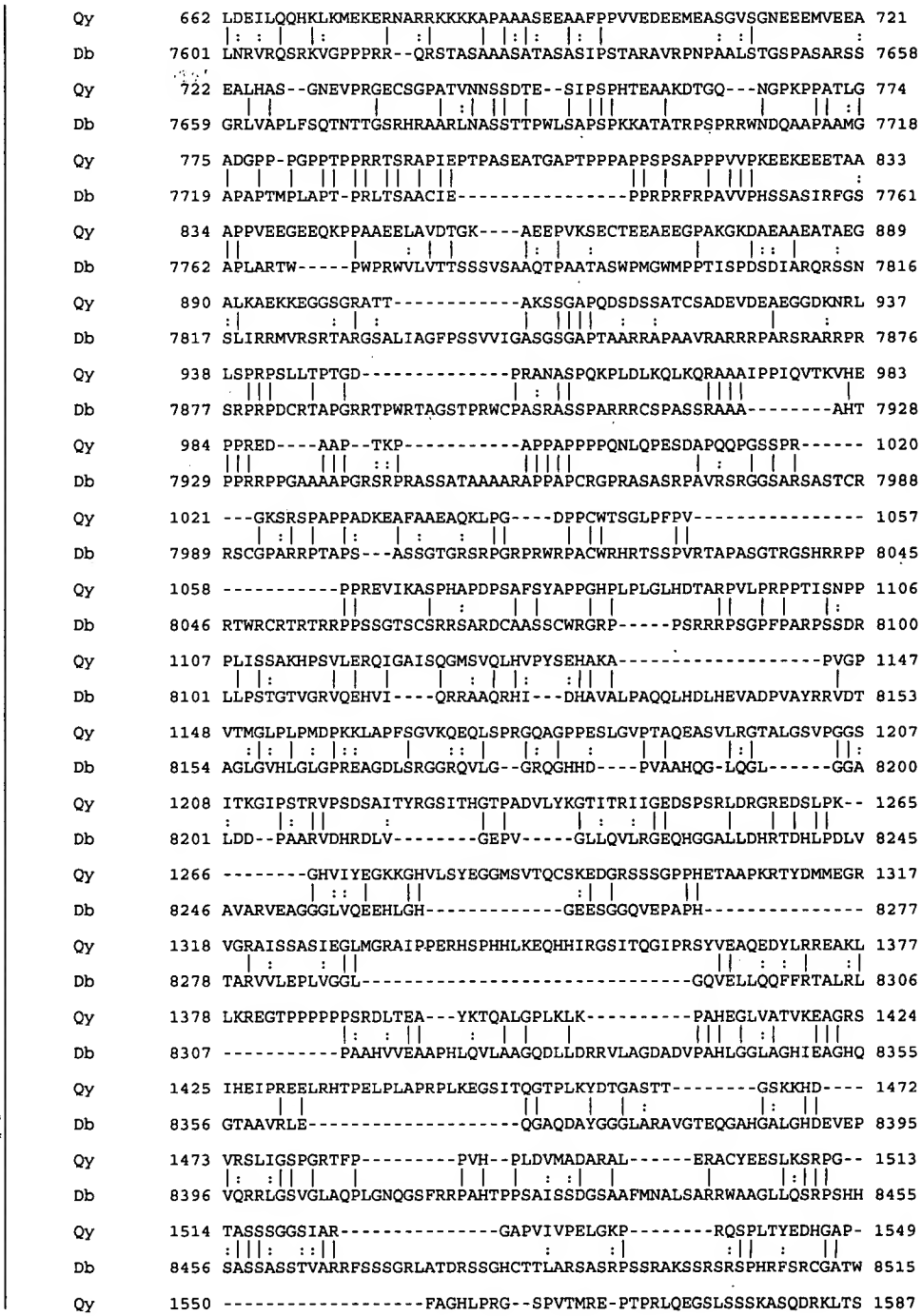
RESULT 71
US-10-210-130-36
; Sequence 36, Application US/10210130
; Publication No. US20040014053A1
; GENERAL INFORMATION:
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Patturajan, Meera
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Miller, Charles E.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Li, Li
; APPLICANT: Berghs, Constance
; APPLICANT: Zhong, Mei
; APPLICANT: Casman, Stacie J.
; APPLICANT: Voss, Edward Z.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Smithson, Glennda
; APPLICANT: Ji, Weizhen
; APPLICANT: Gorman, Linda
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Leite, Mario W.
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Anderson, David W.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Khramtsov, Nikolai V.
; APPLICANT: Ort, Tatiana
; APPLICANT: Ellerman, Karen
; APPLICANT: Rastelli, Luca
; APPLICANT: Agee, Michele L.
; APPLICANT: Chaudhuri, Amitabha
; APPLICANT: Chant, John S.
; APPLICANT: DiPippo, Vincent A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Eisen, Andrew J.
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Giot, Loic
; APPLICANT: Ooi, Chean Eng
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Spaderna, Steven K.
; APPLICANT: Hjalt, Tord
; APPLICANT: Liu, Xiaohong
; APPLICANT: Taupier, Raymond J., Jr.
; APPLICANT: Catterton, Elina
; APPLICANT: Shenoy, Suresh G.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-416C (Cura-716 SMT)
; CURRENT APPLICATION NUMBER: US/10/210,130
; CURRENT FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: 60/309,501
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 60/316,508
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: 60/354,655
; PRIOR FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: 60/310,291
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 60/383,887
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: 60/310,951
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/323,936
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: 60/381,039

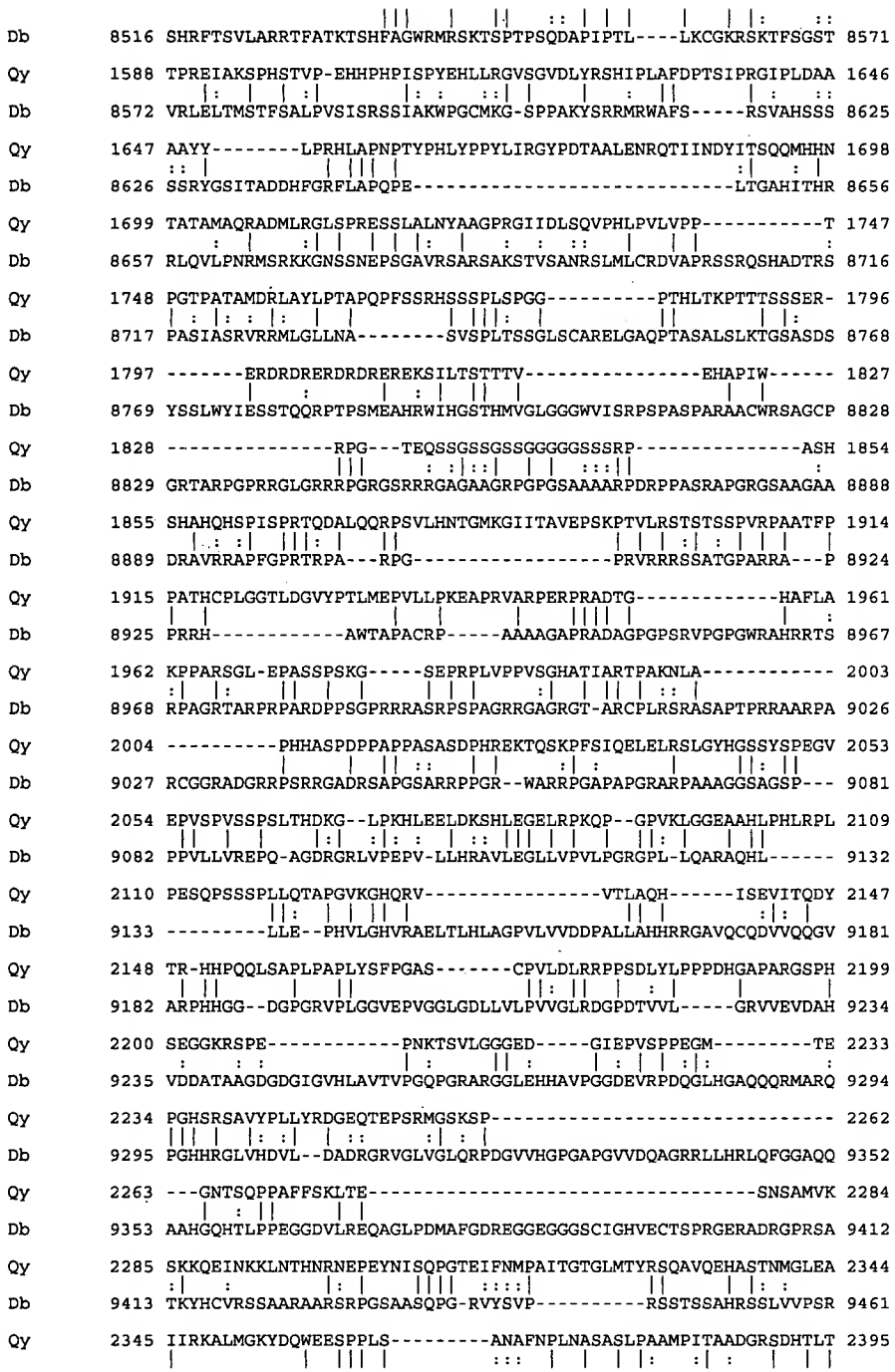
```
;  PRIOR FILING DATE: 2002-05-16
;  PRIOR APPLICATION NUMBER: 60/311,292
;  PRIOR FILING DATE: 2001-08-09
;  PRIOR APPLICATION NUMBER: 60/311,979
;  PRIOR FILING DATE: 2001-08-13
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 369
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 36
;   LENGTH: 3252
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-210-130-36

  Query Match             2.9%;  Score 386;  DB 15;  Length 3252;
  Best Local Similarity   19.0%;  Pred. No. 5.8e-09;
  Matches  478;  Conservative  309;  Mismatches  873;  Indels  860;  Gaps  119;

Qy         78 RSQELHLRPESHSYLPE---LGKSEMEFIESKRPRLELLPDPLLRPSPLLATGQPAGSED 134
              || :: : |:| :|   :   ||  :   : :            ||     |||    | :
Db       1292 RSLDMAIDPDSLTYTVQHQVLGSDQWTALVTG----------LREPGWAATGLRKGVQH 1340

Qy        135 LTKDRSLTGKLEPVSPPSPP-------HTDPELELVPPRLSKEELIQNMDRVDREITMVE 187
              : :  | | |    | ||||          | ||  |  | | :::          :||
Db       1341 IFRVLSTTVKSS--SKPSPPSEPVQLLEHGPTLEEAPAMLDKPDIVY----------VVE 1388

Qy        188 QQISKLKKKQQQLEEEAAKPPEPEKPVSPPPIESKHRSLVQIIYDENRKKAEAAHRILEG 247
              |                    | |           | : :  |  :|:   |   |
Db       1389 GQ----------------------PAS----------VTVTF--NHVEAQVVWRSCRG 1412

Qy        248 LGPQVELPLY--NQPSDTRQYHENIKINQAMRKKL--ILYFKRRNHARKQWKQKFCQRYD 303
                 :  :|  :|| |  ||   ::| :  |: :  :     |  |  :      |
Db       1413 ALLEARAGVYELSQPDDD-QY--CLRICRVSRRDMGALTCTARNRHGTQT-----CSVTL 1464

Qy        304 QLMEA--LEKKVERIENNPRRRAKESKVREYYEKQFPEIR------------------KQ 343
              :| ||   |  :| :|      |: : | |   |  |:|                   ::
Db       1465 ELAEAPRFESIMEDVEVGAGETARFAVVVE--GKPLPDIMWYKDEVLLTESSHVSFVYEE 1522

Qy        344 RELQERMQSRVGQRGSGLSMSAARSEHEVS---EI----------IDGLSEQENLEKQMR 390
              |   : |  | |   : :|    |||   |:          ::|: | |  : : |
Db       1523 NECSLVVLSTGAQDGGVYTCTAQNLAGEVSCKAELAVHSAQTAMEVEGVGEDE--DHRGR 1580

Qy        391 QLAVIPPMLYDADQQ--RIKFINMNGLMADPMKVYKDRQVMNMWSEQEKETFREKFM--- 445
              :|:      ||  |:  |  |  :             |:::   |  |   |  ||:
Db       1581 RLS----DFYDIHQEIGRGAFSYL------------RRIVERSSGLE---FAAKFIPSQ 1620

Qy        446 QHPKNFGLIASFLERKTVAECVLYYYLTKKNENYKSLVRRSYRRRGKSQQQQQQQQQQQQ 505
              ||      : |  :  :||||::     | ::       ||||     :   ::  :
Db       1621 AKPKASARREARLLARLQHDCVLYFH-----EAFE-------RRRGLVIVTELCTEELLE 1668

Qy        506 QQQQQPMPRSSQEEKDEKE--------KEKEAEKEEEKPEVENDKEDLLKEKTDDTSGED 557
              :  ::|   |:      ::        :      : ||      |:||   |  :||
Db       1669 RIARKPTVCESEIRAYMRQVLEGIHYLHQSHVLHLDVKP------ENLL--VWDGAAGE- 1719

Qy        558 NDEKEAVASKGRKTANSQGRRKGR-----------ITRSMANEA--NSEEAITPQQSAEL 604
                ::  :   |    |:|    |             :   : |::  :    | |
Db       1720 --QQVRICDFG----NAQELTPGEPQYCQYGTPEFVAPEIVNQSPVSGVTDIWPVGVVAF 1773

Qy        605 ASMELNESSRWTEEEMETAKKGLLEHGRNWSAIARMVGSKTVSQCKNFYFNYKKRQNLDE 664
              :  |    |  :      :  | : ||::         ::|:        : |  | :
Db       1774 LLSDRNLPVCWGND------RTTLMNIRNYNVAFEETTFLSLSR--------EARGFLIK 1819

Qy        665 ILQQHKLKMEKERN----------------------ARRKKKKAPAAASEEAAFPPVV 700
              :| | :|:   |                      :||: :::  :          |:
Db       1820 VLVQDRLRPTAEETLEHPWFKTQAKGAEVSTDHLKLFLSRRRWQRSQISYKCHLVLRPIP 1879

Qy        701 E--------------DEEMEASGVSGNEEEMVEEAEALHASGNEVPR---GECSGP--AT 741
              |                : |:| : :   || | |  :    |||    | ||   :
Db       1880 ELLRAPPERVWVTMPRRPPPSGGLSSSSDSEEEELEELPS----VPRPLQPEFSGSRVSL 1935

Qy        742 VNNSSDTESIPSPHTEAAKDTGQNGPKPPATLGADGPPPGPPTPPRRTSRAPIEPTPASE 801
               :  :: |:: :| | ||                        |   :   |||      |:
Db       1936 TDIPTEDEALGTPETGAA----------------------TPMDWQEQGRAP------SQ 1967

Qy        802 ATGAPTPPPAPPSPSAPPPVVPKEEKEEETAAAPPVEEGEEQKPPAAEELAVDTGKAE-- 859
                 ||: | | |||              :| ||       || ::  :||       | |
Db       1968 DQEAPS-PEALPSPG------------QEPAAGASPRRGELRRGSSAESALPRAGPRELG 2014

Qy        860 ----EPVKSECTEEAEEGP-----AKGKDAEAAEATAEGALKAEKKEGG--SGRATTAKS 908
                  :    | :    ||     |:|    |   |    ||:      ||   |: : :
Db       2015 RGLHKAASVELPQRRSPGPGATRLARGGLGEGEYAQRLQALRQRLLRGGPEDGKVSGLRG 2074

Qy        909 ------SGAPQDSDSSATCSADEVDEAEGGDKNRLLSPRPSLLTPTGDPRA-NASPQKPL 961
                     |  :|   :   |::     :   :|| |    |       :||  :     ||
Db       2075 PLLESLGGRARDPRMARAASSEAAPHHQPPLENRGLQKSSSFSQGEAEPRGRHRRAGAPL 2134

Qy        962 DLKQLK---QRAAAIPPIQVTKVHEPPREDAAPTKPAPPAPPPPQNLQPE----SDAPQQ 1014
                ::   :   :|    | :       :|    ::| :|: | |  |:: :|     |||||
Db       2135 EIPVARLGARRLQESPSLSALSEAQP----SSPARPSAPKPSTPKSAEPSATTPSDAPQP 2190

Qy       1015 PGSSPRGKSRSPAPPADKEAFAAEAQKLPGDPPCWTSGLPFPVPPR----EVIKASPHAP 1070
                |   | : ::| | :       | | |  ||    |  |: |     : :: | ||
Db       2191 PAPQP-AQDKAPEPRPE----PVRASK-PAPPPQALQTLALPLTPYAQIIQSLQLSGHAQ 2244

Qy       1071 DPSAFSYAPPGHPLPLGLHDT--ARPVLPRP-------PTISNPPPLISSAKHPSVLERQ 1121
                ||    |||  | |   |    ||   | |        |:   || |   |: |:|  |
Db       2245 GPSQGPAAPPSEPKP---HAAVFARVASPPPGAPEKRVPSAGGPPVLAEKARVPTVPPRP 2301

Qy       1122 IGAISQGM-SVQLHVPY-SEHAKAPVGPVTMGLPL--PMDPKKLAPFSGVKQEQLSPRGQ 1177
                ::|  : :::    : ::  ::   |:::|| |      ::  || | ::|
Db       2302 GSSLSSSIENLESEAVFEAKFKRSRESPLSLGLRLLSRSRSEERGPFRGAEEED------ 2355

Qy       1178 AGPPESLGVPTAQEASVLRGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHGTPADV 1237
                  |:      |    || |  |      :: :   |   :  :  | |:
Db       2356 -------GIYRPSPA----GTPLELVRRPERSRSVQDLRAVGEPGLVRRLSL-------- 2396

Qy       1238 LYKGTITRIIGEDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSS 1297
                                |   | |     : |  :| : |   | ||           | |:
Db       2397 --------------SLSQRLRRTPPAQRHPAWEARGGDGESSEG----------GSSA 2430

Qy       1298 SGPPHETAAPKR----TYDMMEGRVGRAISSASIEGLMGRAIPPERHSPHHLKEQHHIRG 1353
                | |     | :|    | : :  |: |: ||    |  ||: |           |   :|
Db       2431 RGSP--VLAMRRRLSFTLERLSSRLQRSGSSEDSGGASGRSTPLFGRLRRATSEGESLR- 2487

Qy       1354 SITQGIPRSYVEAQEDYLRREAKLLKREGT----PPPPPPSRD-----LTEAYKTQALGP 1404
                  |:| : : ||      |: | | |    |  ||       :   | : | |
Db       2488 --RLGLPHNQLAAQAGATTPSAESLGSEASATSGSSAPGESRSRLRWGFSRPRKDKGLSP 2545

Qy       1405 LKLKPAHEGLVATVKEAGRSIHEIPREELRHTPELPLAPRPLKEGSITQGTPLKYDTGAS 1464
                |       :  :| |   |:  | |    |   :    ||:  : :|
Db       2546 PNLS------ASVQEELG---HQYVRSESDFPPVFHI---KLKDQVLLEG---------- 2583

Qy       1465 TTGSKKHDVRSLIGSPGRTFPPVHPLDVMADARALERACYEESLKSRPG---TASSSGGS 1521
                :  :|:  |      |   :  | |          ::||:| |     :   |
Db       2584 -------EAATLLCLPAAC--PAPHISWMKD---------KKSLRSEPSVIIVSCKDGRQ 2625

Qy       1522 IARGAPVIVPELGKPRQSPLTYEDHGAPFAGHLPRGSPVTMREPTPRLQEGSLSSSKASQ 1581
                :    : :|  || | : | ||       | :   |
Db       2626 L-----LSIPRAGK-RHAGL-YECSATNVLGSITSSCTVA------------------- 2658

Qy       1582 DRKLTSTPREIAKSPHSTVPEHHPHPISPYEHLLRGVSGVDLYRSHIPLAFDPTSIPRGI 1641
                        :|: |    |                        :|  :  |::
Db       2659 ----------VARVPGKLAP----------------------PEVPQTYQDTALVLWK 2684

Qy       1642 PLDAAAAYYLPRHLAPNPTYPHLYPPYLIRGYPDTAALENRQTIINDYITSQQMHHNTAT 1701
                | |: |                        | |  | | |       :  : : |  ::
Db       2685 PGDSRA------------------------PCTYTLERR-------VDGESVWHPVSS 2711

Qy       1702 AMAQRADMLRGLSPRESSLALNYAAGPRGIID-LSQVPHLPVLVPPTPGTPATAMDRLAY 1760
                                          || |   | |||| |          |   |:|
Db       2712 ---------------------------GIPDCYYNVTHLPVGV--------TVRFRVAC 2735

Qy       1761 LPTAPQ-PFSSRHSSSPLSPGGPTHLTKPTTTSSSERERDRDRERDRDRDREREKSILTSTT 1819
                  | | |||                    :|||:
Db       2736 ANRAGQGPFS--------------------NSSEK---------------------- 2750
```

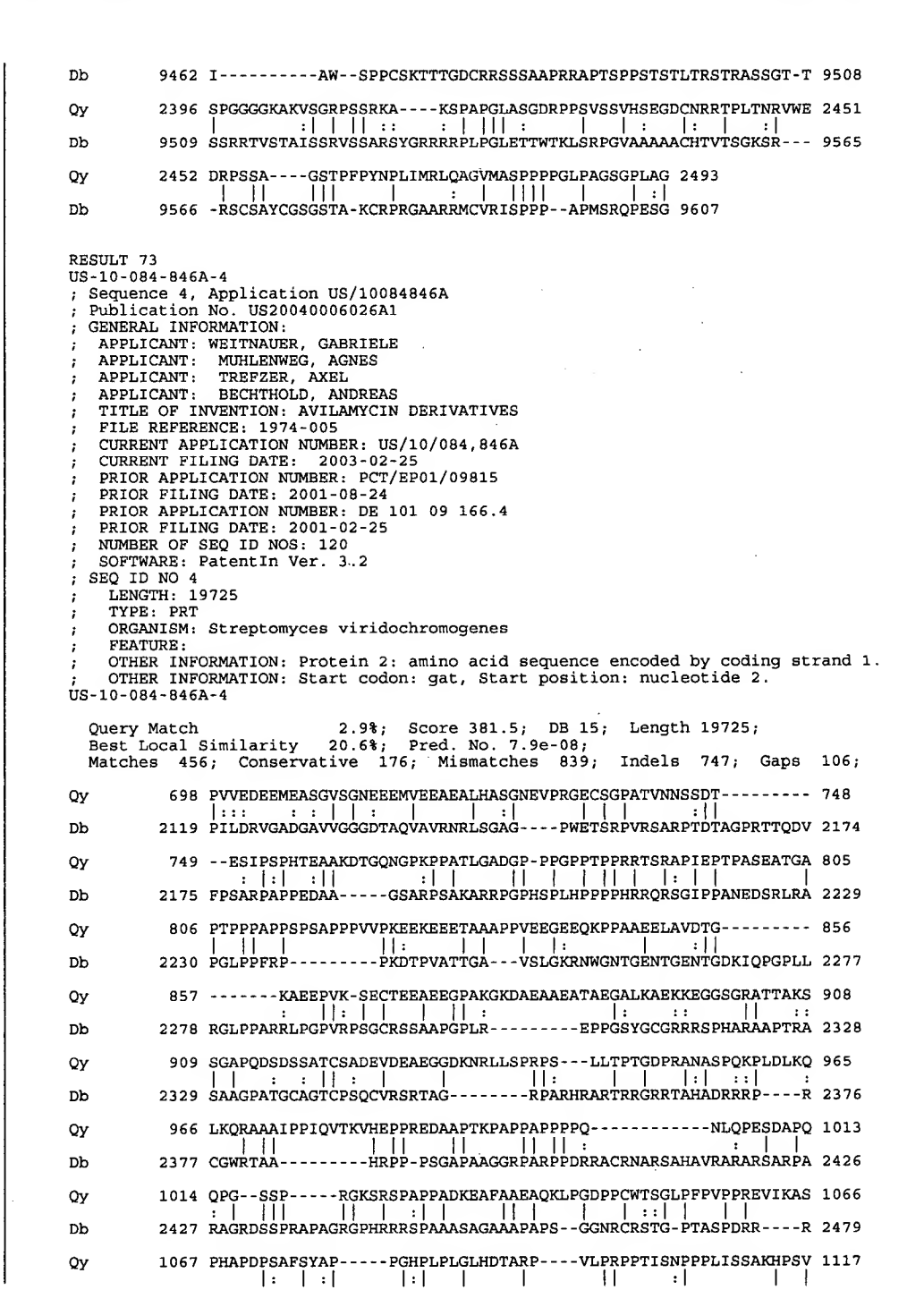


RESULT 72
US-10-084-846A-6
; Sequence 6, Application US/10084846A
; Publication No. US20040006026A1
; GENERAL INFORMATION:
; APPLICANT: WEITNAUER, GABRIELE
; APPLICANT: MUHLENWEG, AGNES
; APPLICANT: TREFZER, AXEL
; APPLICANT: BECHTHOLD, ANDREAS
; TITLE OF INVENTION: AVILAMYCIN DERIVATIVES
; FILE REFERENCE: 1974-005
; CURRENT APPLICATION NUMBER: US/10/084,846A
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: PCT/EP01/09815
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: DE 101 09 166.4
; PRIOR FILING DATE: 2001-02-25
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 6
; LENGTH: 19662
; TYPE: PRT
; ORGANISM: Streptomyces viridochromogenes
; FEATURE:
; OTHER INFORMATION: Protein 1: amino acid sequence encoded by coding strand 2.
; OTHER INFORMATION: Start codon: gga, Start position: nucleotide 1.
US-10-084-846A-6

Query Match 2.9%; Score 383.5; DB 15; Length 19662;
Best Local Similarity 20.1%; Pred. No. 6.4e-08;
Matches 467; Conservative 194; Mismatches 852; Indels 813; Gaps 105;

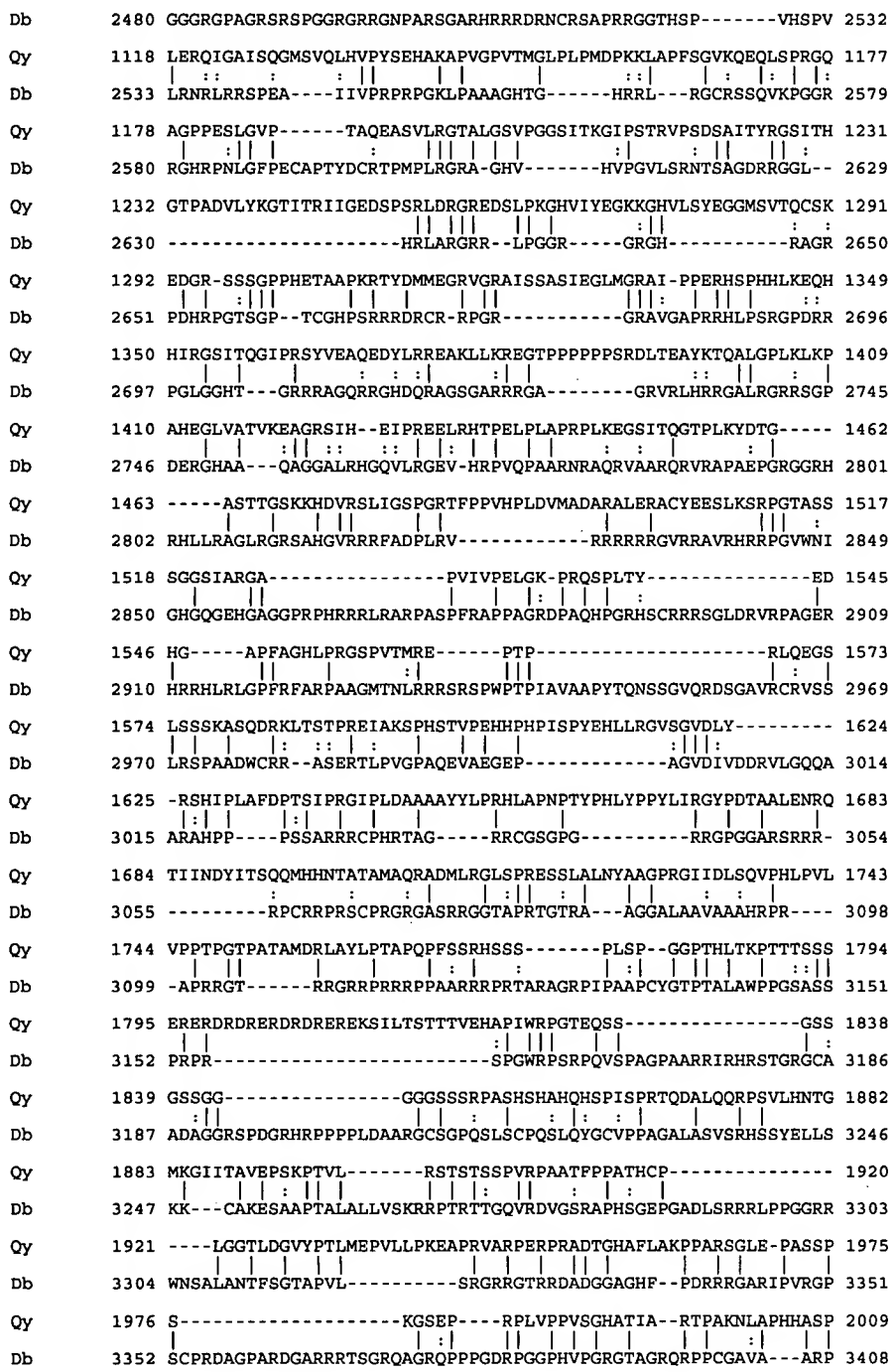


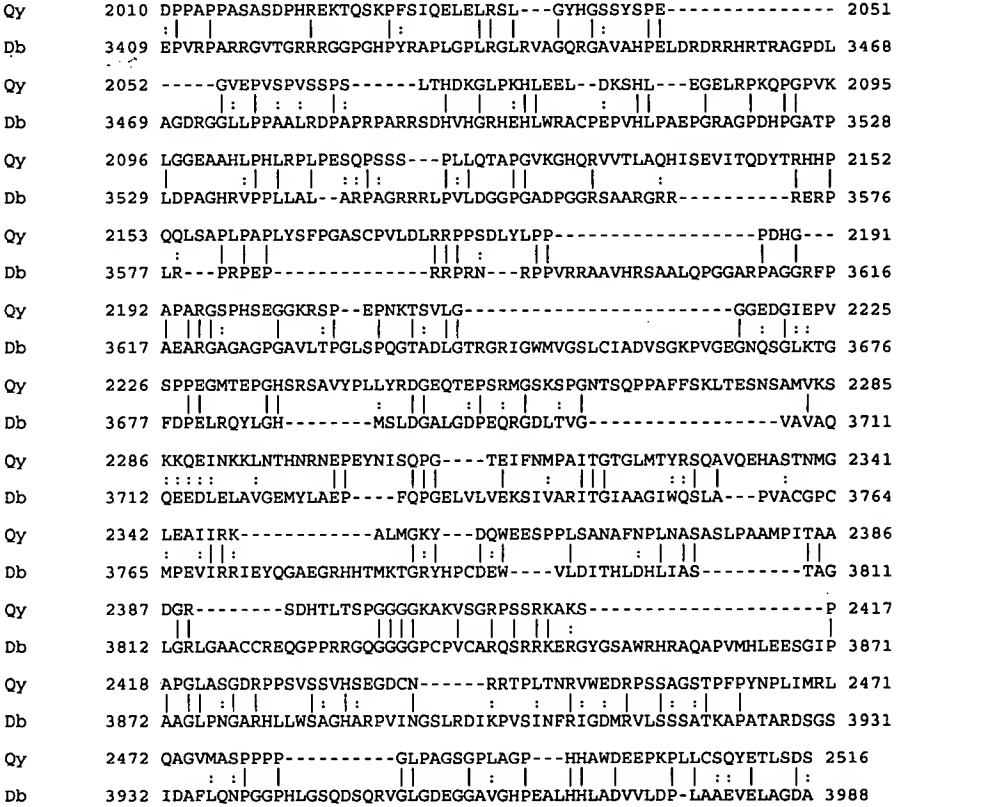




RESULT 73
US-10-084-846A-4
; Sequence 4, Application US/10084846A
; Publication No. US20040006026A1
; GENERAL INFORMATION:
; APPLICANT: WEITNAUER, GABRIELE
; APPLICANT: MUHLENWEG, AGNES
; APPLICANT: TREFZER, AXEL
; APPLICANT: BECHTHOLD, ANDREAS
; TITLE OF INVENTION: AVILAMYCIN DERIVATIVES
; FILE REFERENCE: 1974-005
; CURRENT APPLICATION NUMBER: US/10/084,846A
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: PCT/EP01/09815
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: DE 101 09 166.4
; PRIOR FILING DATE: 2001-02-25
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 4
; LENGTH: 19725
; TYPE: PRT
; ORGANISM: Streptomyces viridochromogenes
; FEATURE:
; OTHER INFORMATION: Protein 2: amino acid sequence encoded by coding strand 1.
; OTHER INFORMATION: Start codon: gat, Start position: nucleotide 2.
US-10-084-846A-4

Query Match 2.9%; Score 381.5; DB 15; Length 19725;
Best Local Similarity 20.6%; Pred. No. 7.9e-08;
Matches 456; Conservative 176; Mismatches 839; Indels 747; Gaps 106;





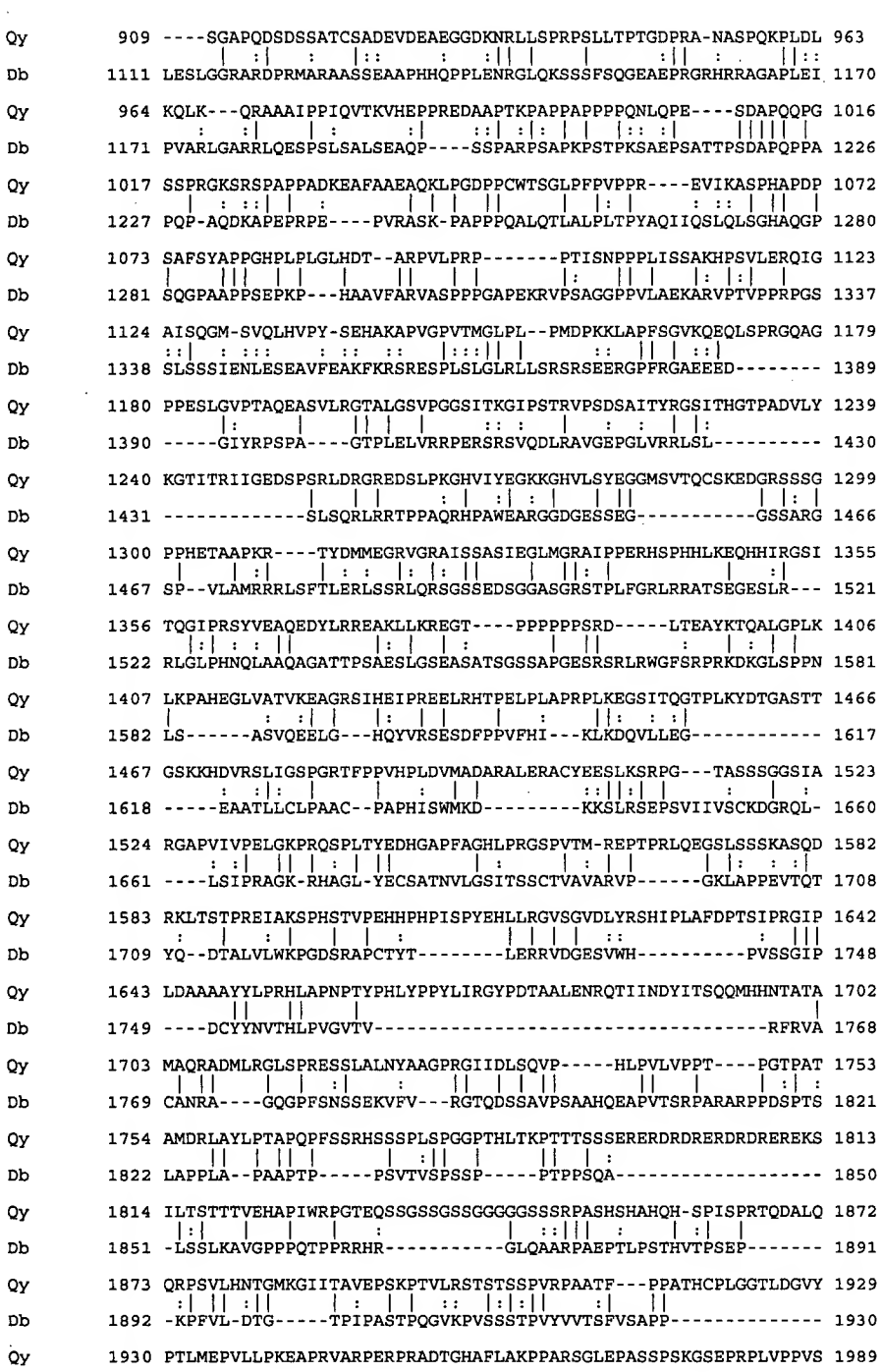
RESULT 74
US-10-359-012-8
; Sequence 8, Application US/10359012
; Publication No. US20030232419A1
; GENERAL INFORMATION:
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
; APPLICANT: KOLODKIN, Alex L.
; APPLICANT: TERMAN, Jon R.
; APPLICANT: MAO, Tianyi
; APPLICANT: PASTERKAMP, Ronald J.
; APPLICANT: YU, Hung-Hsiang
; TITLE OF INVENTION: MOLECULES INTERACTING WITH CASL (MICAL) POLYNUCLEOTIDES, POLYPEPT
; TITLE OF INVENTION: AND METHODS OF USING THE SAME
; FILE REFERENCE: JHU1840-3
; CURRENT APPLICATION NUMBER: US/10/359,012
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: US 60/388,325
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 60/384,302
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: US 60/354,178
; PRIOR FILING DATE: 2002-02-04
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 4723
; TYPE: PRT

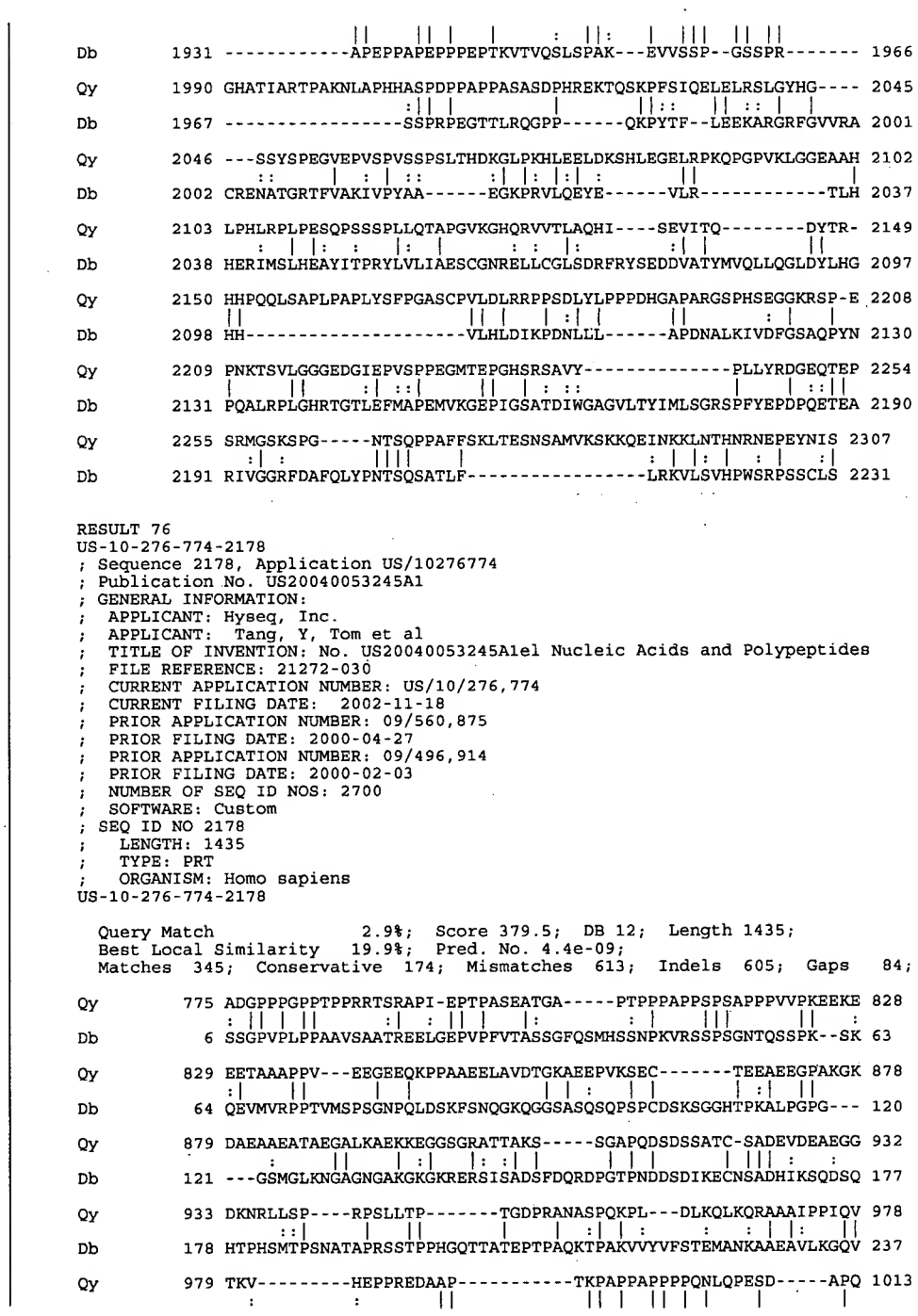
; ORGANISM: Drosophila
US-10-359-012-8

Query Match 2.9%; Score 380.5; DB 15; Length 4723;
Best Local Similarity 19.7%; Pred. No. 1.6e-08;
Matches 574; Conservative 372; Mismatches 1094; Indels 873; Gaps 144;

```
Qy        71 EFQPGNERSQELHLRPESHSYLPE-----LGKSEMEFIESKRPRLELLPDPLLRPSPLLA 125
Db      1834 EYRSPSSEYQRLFQDEKSPSDLSDNIDRILEESELD-VELGLPK---------RSSKKLV 1883

Qy       126 TGQPAGSEDL----TKDRSLTG------KLEPVSPPSPPHTDPE-------LELVPPRL- 167
Db      1884 KTKSLGEGDFDMKPSKER-LTGVQNILKRFESMSSVTSQNSDEQAGFKLRRMESTTSNLS 1942

Qy       168 ----SKEELIQNMDRV-DREITM-------------VEQQISKLKKKQQQLEEEA---A 205
Db      1943 SLTRSRESLVSVSDSMSDLEKTMDYLRNEWRNEATNFLQKKRDKFYAKKEEQEKEAKILA 2002

Qy       206 KPPEPEKPVSPPPIESKHRSLVQII------YDENRK-------------KAEAAHRILE 246
Db      2003 KP----DPLDNLPVQYRDSKLAKFFGLAASKSPENRKSPIKKKKSPSKTPKVTKANNSLE 2058

Qy       247 GLG------------PQVELPLYNQP-SDTRQYHENIKINQAMRKKLILYFKRRNHARKQ 293
Db      2059 ELAKISNVRQTKKAQPKTLKPVEVKPLKPASPVPDDFEILDLLEKATEAKELERSKTKSP 2118

Qy       294 WKQKFCQRYDQLMEALEKKVERIENNPR----------RRAKESKVREYYEKQFPEIRKQ 343
Db      2119 AVESISQTPKEAIVEISLPVEDIKNLPKTGCDKSSNSSRRGSQSSL--IMSRRHSEISLN 2176

Qy       344 RELQERMQSRVGQRGSGLSMSAARSEHEVSEIIDGLSEQENLEKQMRQLAVIPPMLYDAD 403
Db      2177 EKLNQDALAALNQ------IEKEREAEQVDELFQSMVEE--MEQEPQPTAIVEPPEEDID 2228

Qy       404 QQRIKFINMNGLMADPMKVYK-----DRQVMNMWSEQEKETF-REKFMQHPKNFGLI--A 455
Db      2229 ADSLCTTISKSPSAQPVTVVKRGSSEDQSIEKLFSHFSDEMLVNVEFDSNDELVGITPRA 2288

Qy       456 SFLERKTVAECVL--YYYLTKKNENYKSLVRRSY-----------------RRRGKSQQ 495
Db      2289 TLVSRNTEDRDYLDKLESLERDEETFQPVVGEKFIQENVQDEVDGLHFPSRPQRRPKSSS 2348

Qy       496 QQ---------QQQQQQQQQQQQQPMPRSSQ---EEKDEKEKEKEAEK------------ 531
Db      2349 SSSEPSLPVAPQRLEKKLSKLDPEDMPPSVQDLLQQVYQKNIQPELVEVIPVEGKQTLRF 2408

Qy       532 -----EEEKPEVENDKEDLLKEKTDDTSGEDNDEKEAVA-------------SKGRKTAN 573
Db      2409 PSMLAEEDVDEVDHSKEGIKKIETAPEEVRKVTEPEDVARVIPSPIKPSISQSNSLKSEN 2468

Qy       574 SQGRRKGRITRSMANEANSEEAITPQQSAELASMELNESSRWTEEEMETA---KKGLLEH 630
Db      2469 SSGSSLVEIPKIIA---------PPKSSSK------ENSSDWDREKLPASPMPRRRLLPN 2513

Qy       631 GRNWSAIARMVGSKTVSQCKNFYFNYKKRQNLDEILQQHKLK-MEKERNARRKKKKAPAA 689
Db      2514 QTPYKAPS--VASKESS----LEWDMEKLPNSPMLPRRNKMRAISPSTNPVQLLNNLPSD 2567

Qy       690 ASEEAAFPPVVEDEEMEASGVSGNEEEMVEEAEALHASGNEVPRGECSGPATVNNSSDTE 749
Db      2568 VDDEAAQRRLIEDFEQE------RRQALIKRDENFEAIAAEQRRRD--SLQSSSNSSSKR 2619

Qy       750 SIPSP-------------HTEAAKDT-GQNGPKPP--ATLGADGPPPGPPT---PPRRTS 790
Db      2620 SLPPPTPPMMASRRGTTQDTNRTQDTASRHEGTPPMFKKLDVDGSGTSMDSTSCSTRRSS 2679

Qy       791 RAPIE---------PTPASEATGAPTPP---PAPPSPSAPPPVV------PKEEKEEETA 832
Db      2680 FAFIELQDNKPVIVPMPKKLKLPKPEPPRFVPEPVATDEPVPEVFQGRAWPKTQLEGEVD 2739

Qy       833 AAPPVEEGEEQK-------------PPAA--EELAVDTGKAEEPVKSECTEEAE--EG-- 873
Db      2740 LGDSDNEDETEKLKKQLPEYARSDSPPSAAFKNRKWPDGKTVFDKRAESLEEEDIFEGLP 2799

Qy       874 -PAK-----------GKDAEAAEATAEGALKAEK-----KEGGSGRATTAKSSGAPQDSD 916
Db      2800 SPRKRGSQRFMDKPRSQSPQPFKPLANSSRKSSKSFSDLKKGPSLQSLSAQSS---QDTD 2856

Qy       917 SSATCSADEVDEAEGGDKNRLLSPRPSLLTPTGDPRANASPQKPLD-LKQLKQRAAAIPP 975
Db      2857 TLSTTTT-------------VATARPASYANYEDP-MDASTQALLDRSKRLHNRKRDFVN 2902

Qy       976 IQVTKVHEPPREDAAPTKPAPPAPPPPQNLQPESDAPQQPGSSPRGKSRSPAPPADKEAF 1035
Db      2903 ERVVE-RNPYMRDVLRSTDRRDYDDVDEDL--TSYRPRHYASS----------------- 2942

Qy      1036 AAEAQKLPGDPPCWTSGLPFPVPPREVIKASPHAPDPSAFSYAPPGHPLPLGLHDTARPV 1095
Db      2943 --TLNRFPNTTIRKSNNYDYLSPSSDYLSRRSYIPSASATSSYYPSTTRSSHLSDLFRRR 3000

Qy      1096 LPRPPTIS------NPPPLISS-------------AKHPSVLERQIGAISQGMS---VQL 1133
Db      3001 SPASGTVSALSGYGNKESCVISIGLALDRVGHLIESKCTWVRSTKVQTESESTSPDEVEL 3060

Qy      1134 H-------------------VPY----SEHAKAP---------VGPVTMGLPLPMDPKKL 1161
Db      3061 NSATEISTDSEFDNDEIIRQAPKIFIDDTHLRKPTKVQIKSTMIGP------------N 3107

Qy      1162 APFSGVKQEQLSPRGQAG---------PPESLGVPTAQEASVLRGTALGSVPGGSITKGI 1212
Db      3108 AASAGLHQKQLAARE KGGSYLQKYQPQPPLPQFRPLVQVDPTLL--------------I 3152

Qy      1213 PSTRVPSDSAITYRGSITHGTPADVLYKGTITRIIGEDSPSRLDRGREDSLPKGHVIYEG 1272
Db      3153 GSQRAP----------LQNPRPGDYLLNKTAST---EGIASKKSLG----LKKRYLLGEP 3195

Qy      1273 KKGHVLSYEGGMSV-----------TQCSK-----ED---------GRSSSGPPHETAA 1306
Db      3196 ANGNKIQKSGSTSVLDSRIRSFQSNISECQKLLNPSSDISAGMRTFLDRTKLGEGSQT-T 3254

Qy      1307 PKRTYDMMEGRVGRAISSASIEGLMGRAIPPERHSPHHLKEQHHI--RGSITQGIPRSYV 1364
Db      3255 PGQTNELIRSATSNVINDLRVELRIQKT--GSSHSTDNEKENVFVNCKNELNKGM--EYT 3310

Qy      1365 EAQE----DYLRREAKLLKREG----------TPPPPPPSRDLT--EAYKTQALGPLKLK 1408
Db      3311 DAVNATLLDQLARKSSPTTPTNKTVVEVIDLVTPEKPIDIIDLTALETPKKQLV------ 3364

Qy      1409 PAHEGLVATVKEA----GRSIHEIPREELRHTPELPLAPRPLKE------GSITQGTPLK 1458
Db      3365 ---DGSAMDVDERLTPDSNKISEL-QQEVKEEPK-PDVSRDVKECIPDILGHIKEG---- 3415

Qy      1459 YDTGASTTGSKK---HDVRSLIGSPGRTFPPVHPLDVMADARALERACYE-ESL------ 1508
Db      3416 -------TGSKEPGGEDQQSLLEQSDEEKRDSPEKDV------AEHELYEPDSVQIQVPN 3462

Qy      1509 ----KSRPGTASSSG--GSIARGAPVIVPELGKPRQSPLTYEDHGAPFAGHLPRGSPVTM 1562
Db      3463 IPWEKSKPEVMSTTGSSGSICSSSD----------SSSIEDIQHYILESTTSPDTQTVGG 3512

Qy      1563 REPTPRLQEGSLSSSKASQDRKLTSTPREIAKSPHSTVPEHHPHPISPYEHLLRGVSGVD 1622
Db      3513 KHNVPRLEVHDTSGALMQVDSLMIVNGKYIGDPEDVKFLDMPANVIVPPAPALK-TNELD 3571

Qy      1623 LYRSHIPLAFDPTSIPRGIPLDAAAAYYLPRHLAPNPTYPHLYPPYLIRGYPDTAALENR 1682
Db      3572 MEDDQEAEAEPVTATPEPVECTVIEA---ERRVTAPPPLPEMGPPKLKFDSKNENKIESL 3628

Qy      1683 QTIINDYITSQQMHHNTATAMAQRADMLRGLSPRESSLALNYAAGPRGIIDLSQVPHLPV 1742
Db      3629 KNL--PLIVESNVEHSQA---------VKPITLNLSNLA---------------------- 3656

Qy      1743 LVPPTPGTPATAMDRLAYLPTAPQPFSSRHSSSPLSPGGPTHLTKPTTTSS-------SE 1795
Db      3657 RTPDTPTTP-TAHDS-DKTPTG--EILSRGSDSETEHTGTGQVLTETELSDWTADDCISE 3712
```

```
Qy       1796 RERDRD-----------RERDRDREREKSILTSTTTVEHAPIWRPGTEQSSG-------- 1836
Db       3713 NFVDLEFALNSNKGTIKRRKDR-RRSGASKLPSGNEVIHELARQAPVVQMDGILSAIDID 3771

Qy       1837 -----SSGSSGGGGGSSSRPASHSHAHQHSPISPRTQDALQQRPSVLHNTGMKGIITAVE 1891
Db       3772 DIEFMDTGSEG----------SCAEAY------PATNTAL------IQNRGYMEYIEA-E 3808

Qy       1892 PSKPTVLRSTSTSSPVRPAATFPPATHCPLGGTLDGVYPTLMEPVLLPKEAPRVARPERP 1951
Db       3809 PKKTT----------RKAA--PPSS----------YPGNLPPLM-------TKRDEKL 3837

Qy       1952 RAD--TGHAFL----AKPPARSGLEPASSPSKG--------------SEPRPLVPPVSG 1990
Db       3838 GVDYIEQGAYIMHDDAKTPVNE-VAPAMTQSLTDSITLNELDDDSMIISQTQPTTTEESE 3896

Qy       1991 HATIARTPAKNLAPH---------HASPDPPAPPASASDPHREKTQSKPFSIQELELRSL 2041
Db       3897 ALTVVTSPLDTSSPRVLDQFASMLAAGKGDSTPSSSEQQP---KTSTVTSS-------ST 3946

Qy       2042 GYHGSSYSPEGVEPVSPVSSPSLTHDKGLPKHLEELD---KSHLEGE---LRPKQPGPVK 2095
Db       3947 GPNSSTTGNVSKEPQEEDLQIQFEYVRALQQRISQISTQRRKSSKGEAPNLQLNSSAPVI 4006

Qy       2096 LGGEAAHLPHLRPLPESQPSSSPLLQTAPGVKGHQRVVTLAQHISEVITQDYTR------ 2149
Db       4007 ESAEDPAKPAEEPLVSMRPRTTSISGKVP------EIPTLSSKLEE-ITKERTKQKDLIH 4059

Qy       2150 ------------------HHPQQLSAPLPAPLYSFPGASCPVLDLRRPPSDLYLPPPDH 2190
Db       4060 DLVMDKLQSKKQLNAEKRLHRSRQRS--LLTSGYASGSSLSPTPKLAAACS-----PQDS 4112

Qy       2191 GAPARGSPHSEGGKRSPEPNKTSVLGGGEDGIEPVSPPEGMTEPGHSRSAVYPLLYRDGE 2250
Db       4113 NCSSQAHYHASTAEEAPKP----------------PAERPLQKSATSTYVSP--YRTVQ 4153

Qy       2251 QTEPSRMGSKSPGNTSQPPAFFSKLTESN--SAMVKSKKQEIN-KKLNTHNRNEPEYNIS 2307
Db       4154 --APTRSAD------LYKPRPFSEHIDSNALAGYKLGKTASFNGGKLGDFAKPIAPARVN 4205

Qy       2308 QPG----TEIFNMPAITGTGLMTYRSQA---VQEHASTNMGLE-----AIIRKALMGKYD 2355
Db       4206 RGGGVATADIANISAST----ENLRSEARARARLKSNTELGLSPEEKMQLIRSRL--HYD 4259

Qy       2356 --------QWEESP--PLSANAFNPLNASASL----------------PAAMPITAADG 2388
Db       4260 QNRSLKPKQLEEMPSGDLAARA-RKMSASKSVNDLAYMVGQQQQQQVEKDAVLQAKAADF 4318

Qy       2389 RSDHTLTSPGGGGKAKVSGRPSSRKAKSP----------------APGLASGDRPPS 2429
Db       4319 TSDPNLAS---GGQEKAGKTKSGRRPKDPERRKSLIQSLSSFFQKGSGSAASSSKEQGGA 4375

Qy       2430 VSSVHSEGDCNRRTPLTNRVWEDRP-SSAGSTP 2461
Db       4376 VAAVHSE------------QSERPGTSSSGTP 4395


RESULT 75
US-10-379-381-5
; Sequence 5, Application US/10379381
; Publication No. US20030228595A1
; GENERAL INFORMATION:
;  APPLICANT: GAN, Weiniu et al.
;  TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
;  TITLE OF INVENTION:  ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: CL001360
;  CURRENT APPLICATION NUMBER: US/10/379,381
;  CURRENT FILING DATE:  2003-03-05
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;  LENGTH: 2231
;  TYPE: PRT
;  ORGANISM: Homo sapiens
US-10-379-381-5

  Query Match             2.9%;  Score 380;  DB 15;  Length 2231;
  Best Local Similarity   19.7%;  Pred. No. 7e-09;
  Matches  485;  Conservative  296;  Mismatches  897;  Indels  780;  Gaps  120;

Qy         78 RSQELHLRPESHSYLPE---LGKSEMEFIESKRPRLELLPDPLLRPSPLLATGQPAGSED 134
Db        326 RSLDMAIDPDSLTYTVQHQVLGSDQWTALVTG----------LREPGWAATGLRKGVQH 374

Qy        135 LTKDRSLTGKLEPVSPPSPP-------HTDPELELVPPRLSKEELIQNMDRVDREITMVE 187
Db        375 IFRVLSTTVKSS--SKPSPPSEPVQLLEHGPTLEEAPAMLDKPDIVY----------VVE 422

Qy        188 QQISKLKKKQQQLEEEAAKPPEPEKPVSPPPIESKHRSLVQIIYDENRKKAEAAHRILEG 247
Db        423 GQ----------------------PAS----------VTVTF--NHVEAQVVWRSCRG 446

Qy        248 LGPQVELPLY--NQPSDTRQYHENIKINQAMRKKL--ILYFKRRNHARKQWKQKFCQRYD 303
Db        447 ALLEARAGVYELSQPDDD-QY--CLRICRVSRRDMGALTCTARNRHGTQT-----CSVTL 498

Qy        304 QLMEA--LEKKVERIENNPRRRAKESKVREYYEKQFPEIR----------------KQ 343
Db        499 ELAEAPRFESIMEDVEVGAGETARFAVVVE--GKPLPDIMWYKDEVLLTESSHVSFVYEE 556

Qy        344 RELQERMQSRVGQRGSGLSMSAARSEHEVS---EI----------IDGLSEQENLEKQMR 390
Db        557 NECSLVVLSTGAQDGGVYTCTAQNLAGEVSCKAELAVHSAQTAMEVEGVGEDE--DHRGR 614

Qy        391 QLAVIPPMLYDADQQ--RIKFINMNGLMADPMKVYKDRQVMNMWSEQEKETFREKFM--- 445
Db        615 RLS----DFYDIHQEIGRGAFSYL----------RRIVERSSGLE---FAAKFIPSQ 654

Qy        446 QHPKNFGLIASFLERKTVAECVLYYYLTKKNENYKSLVRRSYRRRGKSQQQQQQQQQQQQ 505
Db        655 AKPKASARREARLLARLQHDCVLYFH-----EAFE-------RRRGLVIVTELCTEELLE 702

Qy        506 QQQQQPMPRSSQEEKDEKE--------KEKEAEKEEEKPEVENDKEDLLKEKTDDTSGED 557
Db        703 RIARKPTVCESEIRAYMRQVLEGIHYLHQSHVLHLDVKP------ENLL--VWDGAAGE- 753

Qy        558 NDEKEAVASKGRKTANSQGRRKGR----------ITRSMANEANSEEAITPQQSAELAS 606
Db        754 --QQVRICDFG----NAQELTPGEPQYCQYGTPEFVAPEIVNQSPVSGVTDIWPVGVVAF 807

Qy        607 MELNESSRWTEEEMETAKKGLLEHGRNWSAIARMVGSKTVSQCKNFYFNYKKRQNLDEIL 666
Db        808 LCLTGISPFVGENDRTT----LMNIRNYNVAFEETTFLSLSR--------EARGFLIKVL 855

Qy        667 QQHKLKMEKERN----------------------ARRKKKKAPAAASEEAAFPPVVE- 701
Db        856 VQDRLRPTAEETLEHPWFKTQAKGAEVSTDHLKLFLSRRRWQRSQISYKCHLVLRPIPEL 915

Qy        702 ------------DEEMEASGVSGNEEEMVEEAEALHASGNEVPR---GECSGP--ATVN 743
Db        916 LRAPPERVWVTMPRRPPPSGGLSSSSDSEEEELEELPS----VPRPLQPEFSGSRVSLTD 971

Qy        744 NSSDTESIPSPHTEAAKDTGQNGPKPPATLGADGPPPGPPTPPRRTSRAPIEPTPASEAT 803
Db        972 IPTEDEALGTPETGAA----------------------TPMDWQEQGRAP------SQDQ 1003

Qy        804 GAPTPPPAPPSPSAPPPVVPKEEKEEETAAAPPVEEGEEQKPPAAEELAVDTGKAE---- 859
Db       1004 EAPS-PEALPSPG------------QEPAAGASPRRGELRRGSSAESALPRAGPRELGRG 1050

Qy        860 --EPVKSECTEEAEEGP-----AKGKDAEAAEATAEGALKAEKKEGG--SGRATTAKS-- 908
Db       1051 LHKAASVELPQRRSPGPGATRLARGGLGEGEYAQRLQALRQRLLRGGPEDGKVSGLRGPL 1110
```





RESULT 76
US-10-276-774-2178
; Sequence 2178, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y, Tom et al
; TITLE OF INVENTION: No. US20040053245A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: US/10/276,774
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 2178
; LENGTH: 1435
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-276-774-2178

Query Match 2.9%; Score 379.5; DB 12; Length 1435;
Best Local Similarity 19.9%; Pred. No. 4.4e-09;
Matches 345; Conservative 174; Mismatches 613; Indels 605; Gaps 84;

```
Db        238 ETIVSFHIQNISNNKTERSTAPLNTQISALRNDPKPLPQQPPAPANQDQNSSQNTRLQPT 297

Qy       1014 QPGSSPRGKSRSPAPPADKEAFAAEAQKLP--GDPPCWTSGLPFPVPP----------R 1060

Db        298 PPIPAPAPKPAAPPRPLDRESPGVENKLIPSVGSPASST-----PLPPDGTGPNSTPNNR 352

Qy       1061 EVIKASPHAPDPSAFSYAPPGHPLPLGLHDTARPVLPRPPTISNPPPLISSAKHPSVLER 1120

Db        353 AVTPVSQGSNSSSADPKAPPPPPVSSG----------EPPTLGENPDGLSQEQ----LEH 398

Qy       1121 QIGAISQGMSVQLHVPYSEHAKAPVGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGP 1180

Db        399 RERSLQTLRDIQ---------------RMLFP---DEKE---FTGAQ--------SGGP 428

Qy       1181 PESLGV---PTAQ-----EASVLRGTALGSVPGGSITKGI-------------PSTRVP- 1218

Db        429 QQNPGVLDGPQKKPEGPIQAMMAQSQSLGKGPGPRTDVGAPFGPQGHRDVPFSPDEMVPP 488

Qy       1219 ---SDSAITYRGSITHGTPADVLYKGTITRIIGEDSPSRLDRGREDSLPKGHVIYEGKKG 1275

Db        489 SMNSQSGTIGPDHLDHMTPEQIAW----------------------LKLQQEFYEEKR- 524

Qy       1276 HVLSYEGGMSVTQCSKEDGR-SSSGPPHETAAPKRTYDM--MEGRVGRAISSASIEGL-M 1331

Db        525 ---RKPEQVVVQQCSLQDMMVHQHGPRGVVRGPPPPYQMTPSEGWAPGGTEPFS-DGINM 580

Qy       1332 GRAIPPERHSPHHLKEQHHIRGSITQGIPRSYVEAQEDYLRREAKLLKREGTPPPPPPSR 1391

Db        581 PHSLPPRGMAPH-----------------------------------------PNMPGSQ 599

Qy       1392 DLTEAYKTQALGPLKLKPAHEGLVATVKEAGRSIHEIPREELRHTPELPLAPRPLKEGSI 1451

Db        600 ------------MRL-PGFAGMINSEMEG-----------------PNVPNPASRPGL 627

Qy       1452 TQGTPLKYDTGASTTGSKKHDVRSLIGSPGR--TFPPVHPLDVMADARALERACYEESLK 1509

Db        628 S-GVSWPDDVPKIPDGRNFPPGQGIFSGPGRGERFP---------NPQGLSEEMFQQQLA 677

Qy       1510 SR-----PGTA--------------SSSGGSIARGAPVIVPELGKPRQSPLTYEDHGAPF 1550

Db        678 EKQLGLPPGMAMEGIRPSMEMNRMIPGSQRHMEPGNNPIFPRI--PVEGPL------SPS 729

Qy       1551 AGHLPRGSPVTM---RE------PTPRLQEGSLSSSKASQDRKLTSTPREIAKSPHSTVP 1601

Db        730 RGDFPKGIPPQMGPGRELEFGMVPSGMKGDVNLNVNMGSNSQMIPQKMREAGAGP----- 784

Qy       1602 EHHPHPISPYEHLLR----GVSGVDLYRSHIPLAFDPTSIPRGIPLDAAAAYYLPRH--- 1654

Db        785 ----------EEMLKLRPGGSDMLPAQQKMVPLPFGEH----------------PQQEYG 818

Qy       1655 LAPNPTYPHLYPPYLIRGYPDTAALEN-RQTIINDYITSQQMHH---------NTATAMA 1704

Db        819 MGPRPFLP------MSQGPGSNSGLRNLREPIGPDQRTNSRLSHMPPLPLNPSSNPTSLN 872

Qy       1705 QRADMLRGL--SPRESSLALNYAAGPRGIIDL-SQVPH---LPVLVPP-----TPGTPAT 1753

Db        873 TAPPVQRGLGRKPLDISVAGSQVHSP-GINPLKSPTMHQVQSPMLGSPSGNLKSPQTPS- 930

Qy       1754 AMDRLAYLPTAPQPFSSRHSSSPL--SPGGPTHLTKPTTTSSSERERDRDRERDRDRERE 1811

Db        931 ---QLAGMLAGPAAAASIKSPPVLGSAAASPVHLKSPSLPAPS----------------- 970

Qy       1812 KSILTSTTTVEHAPIWRPGTEQSSGSSGSSGGGGGSSSRPASHSHAHQHSPISPRTQDAL 1871

Db        971 ------------PGWTSSPEPPLQSPGI----------PPNHK--------APLTM--- 996

Qy       1872 QQRPSVLHNTGMKGIITAVEPSKPTVLRSTSTSSPVRPAATFPPATHCPLGGTLDGVYPT 1931

Db        997 --------------------ASPAMLGNVESGGPPPPTASQPASVNIP--GSLPSSTPY 1033

Qy       1932 LMEPVLLPKEAP------RVARPERPRADTGHAFLA----------KPPARSGLEPA--S 1973

Db       1034 TMPPEPTLSQNPLSIMMSRMSKFAMPSSNPGYNHDAIKTVASSDDDSPPARSPNLPSMNN 1093

Qy       1974 SPSKG--------SEPRPLVP-----------PVSGHATIARTP---AKNLAPH------ 2005

Db       1094 MPGMGINTQNPRISGPNPVVPMPTLSPMGMTQPLS-HSNQMPSPNAVGPNIPPHGVPMGP 1152

Qy       2006 ----------HASPDPP----------------APPASASDPHREKT------------ 2026

Db       1153 GLMSHNPIMGHGSQEPPMVPQGRMGFPQGFPPVQSPPQQVPFPHNGPSGGQGSFPGGMGF 1212

Qy       2027 -------------------------------------QSKPFSIQELELRSLGYHGSSYS 2049

Db       1213 PGEGPLGRPSNLPQSSADAALCKPGGPGGPDSFTVLGNSMPSVFTDPDLQEVIRPGATGI 1272

Qy       2050 PE-GVEPVSPVSSPSLTHDKGLPKHLEELDKSHLEGELRPKQPGP-----VKLGGEAA-- 2101

Db       1273 PEFDLSRIIPSEKPSQT--------LQYFPRGEVPGRKQPQGPGPGFSHMQGMMGEQAPR 1324

Qy       2102 ---HLPHL-RPLPESQPSSSPLLQTAPGVKGHQRVVTLAQHISEVITQDYTRHHPQQLSA 2157

Db       1325 MGLALPGMGGPGPGVGTPDIP--LGTAPSMPGHNPMRPPAFLQQGMMGPHHRMMSPAQSTM 1382

Qy       2158 PLPAPLYSFPGASCPVLDLR-RPPSDLYLPPPDHGAPARGSPHSEGGKRSPEPNKTS 2213

Db       1383 PGQPTLMSNPAAAVGMIPGKDRGPAGLYTHPGPVGSP--GMMMSMQGMMG--PNRTS 1435
```

RESULT 77
US-10-051-874-169
; Sequence 169, Application US/10051874
; Publication No. US20040005557A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Alsobrook II, John P
; APPLICANT: Colman, Steven D
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Boldog, Ferenc
; APPLICANT: Vernet, Corine AM
; APPLICANT: Li, Li
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Casman, Stacie J
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Edinger, Shlomit R
; APPLICANT: MacDougall, John R
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard A
; APPLICANT: Pena, Carol EA
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Millet, Isabelle
; APPLICANT: Miller, Charles E
; APPLICANT: Lepley, Denise M
; APPLICANT: Smithson, Glennda
; APPLICANT: Baumgartner, Jason C
; APPLICANT: Herrman, John L
; APPLICANT: Peyman, John A
; APPLICANT: Gorman, Linda
; APPLICANT: Mezes, Peter D
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Gerlach, Valerie
; APPLICANT: Grosse, William M
; APPLICANT: Liu, Xiaohong
; APPLICANT: Ellerman, Karen
; APPLICANT: Rothenberg, Mark
; APPLICANT: Stone, David J
; APPLICANT: Burgess, Catherine E
; TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
; TITLE OF INVENTION: USING THE SAME
; FILE REFERENCE: 21402-245
; CURRENT APPLICATION NUMBER: US/10/051,874
; CURRENT FILING DATE: 2002-09-25
; PRIOR APPLICATION NUMBER: 60/268,595

```
;  PRIOR FILING DATE: 2001-02-14
;  PRIOR APPLICATION NUMBER: 60/325,306
;  PRIOR FILING DATE: 2001-09-27
;  PRIOR APPLICATION NUMBER: 60/262,587
;  PRIOR FILING DATE: 2001-01-18
;  PRIOR APPLICATION NUMBER: 60/272,409
;  PRIOR FILING DATE: 2001-02-28
;  PRIOR APPLICATION NUMBER: 60/262,454
;  PRIOR FILING DATE: 2001-01-18
;  PRIOR APPLICATION NUMBER: 60/276,777
;  PRIOR FILING DATE: 2001-03-16
;  PRIOR APPLICATION NUMBER: 60/291,672
;  PRIOR FILING DATE: 2001-05-17
;  PRIOR APPLICATION NUMBER: 60/330,336
;  PRIOR FILING DATE: 2001-10-18
;  PRIOR APPLICATION NUMBER: 60/265,530
;  PRIOR FILING DATE: 2001-01-31
;  PRIOR APPLICATION NUMBER: 60/261,376
;  PRIOR FILING DATE: 2001-01-16
;  NUMBER OF SEQ ID NOS: 269
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 169
;   LENGTH: 4823
;   TYPE: PRT
;   ORGANISM: Takifugu rubripes
US-10-051-874-169

  Query Match             2.9%;  Score 377.5;  DB 15;  Length 4823;
  Best Local Similarity   18.4%;  Pred. No. 2.3e-08;
  Matches  542;  Conservative  345;  Mismatches 1146;  Indels  917;  Gaps  126;

Qy          53 LSPGS-----IIQPQRRRPSLLSEFQPGNE-RSQELHLRPESHSYLPELGKSEMEFI--- 103
               : ||       :::|    || |     |  | |:| :    ||      |: :|   |
Db        1104 IPPGQRPASCLVRPSLPAPSGLRITSPECESRNQSIFFILESQKPYCEVTQSFFFFFAAD 1163

Qy         104 ESKRPRLELLP-------------DPLLRPSPLLATGQPAGSEDL-----------TKDR 139
                 |    :: |             |          || | | |:             |
Db        1164 ASNHVAKDIFPLKQEGCEQSQAQKD---------GTGLPQGVENQDSEQFFRKVLGVSDG 1214

Qy         140 SLTGKLEPV--SPPSPPHTD-PELELVPPRLSKEELIQNMDRVDREITMVEQQISKLKKK 196
               |  | ::|:  |      || :|:  |:|  |   |::     : :  :  :    |
Db        1215 SSLGGMKPILESSKGESHTALPQSALLPGSLPSAEMVDAFPGLSQSPFLDMRDRGGLFSP 1274

Qy         197 QQQLEEEAAKPPEPEKPVSPP-PIESKHRSLVQIIYDENRKKAEAAHRILEGLGPQVELP 255
                 |   | |  |  | ||| | |::                    :||
Db        1275 DGGEESPWATPSTPVTPSSPPTPTETEG----------------------DGLS------ 1306

Qy         256 LYNQPSDTRQYHENIKINQAMRKKLILYFKRRNHARKQWKQKFCQRYDQLMEALEKKVER 315
                |||  |  ::: :: :: :      :||     .|:   :  |  |:|: : :||
Db        1307 -YNQRS-LQRWEKDEELGELSTISPVLYANTNFPTLKRDYPDWASRCKQIMK-IWRKVSA 1363

Qy         316 IENNP-RRRAKESKVREYYEKQFPEIRKQRELQ-----ERMQSRVGQRGSGLSMSAARSE 369
                :  |  ::||::  :   |     :|| | |     :    || :|   | :
Db        1364 ADKVPYLQKAKDNRAAQRISK----AQKQAESQVCRPIKTEPGRVKER-PNLHLKIPLPA 1418

Qy         370 HEVSEIIDGLSEQENLEKQMRQLAVIPPMLYDADQQRIKFINMNGLMADPMKVYKDRQVM 429
                ||       |:  : |          |:| |:    : |                | |
Db        1419 GSVS----ASSQPSSAESPF-------PLLPDSGSSSVFF-------------SDGPVR 1453

Qy         430 NMWSEQEKETFREKF-------------MQHPKNFGLIASF------------------ 457
                  | : :     ||             |     | |||
Db        1454 TPGSAEIRTDPLAKFPPQSPHCHSHPPTPFSHAGASPLQASFSGYVPSGPQGPPQGRPAS 1513

Qy         458 -----LERKTV-----AECVLYYYLTKKNENYKSLVRRSYRRRGKSQQQQQQQQQQQQQQ 507
                    ::  |      |: |  |:                | : |:||   | ||
Db        1514 LGPFDMQPGTPGTPRRAQQVDPYF-----------------RSQLQKQQGHLPQTQQGS 1555

Qy         508 QQQPMPRSSQEEKDEKEKEKEAEKEEEKPEVENDKEDLLKEKTDDTSGEDNDEKEAVASK 567
               |:   |  |  :                |    |  |   :  |   |    :
Db        1556 QESLAPPGSPHSR----------------VAGIGESPLFSPSHSTHYGDAFRNQ--QGM 1596

Qy         568 GRKTANSQGRRKGRITRSMANEANSEEAITPQQSAELASMELNESSRWTEEEMETAKKGL 627
               ||    |     |:|         |  | | | |  |:: |:    ||    :    :  |:
```

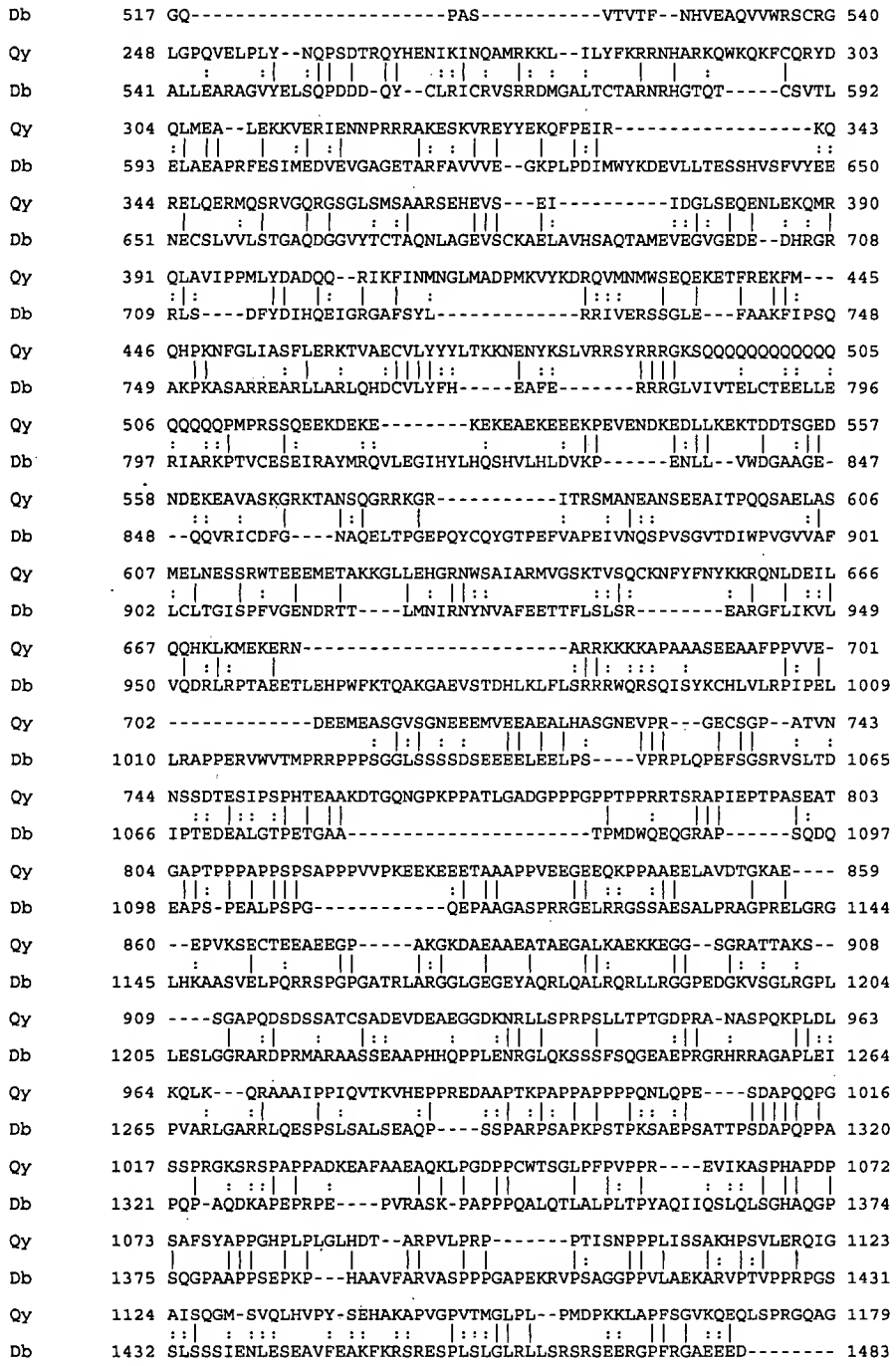
```
Db        1597 GRPEYGSSPSHSGQI---------SSPASTGQYRADM-SVPSPRSSTGRTDLSTGSPAGM 1646

Qy         628 LEHGRNWSAIARMVGSKTVSQCKNFYFNYKKRQNLDEILQQHKL-KMEKERNARRKKKKA 686
               || |                                  | :  :     :| :        :
Db        1647 LESG-------------------------------DGLFKAPMTPRMHQGDGGALHPGAS 1675

Qy         687 PAAASE------EAAFP--PVVEDEEMEASGVSGNEEEMVEEAEALHASGNEVPRGECSG 738
               |:  ||          ||  |::   :   :   | :   : : |          ||    |
Db        1676 PSHPSEGYKQSPSHPFPESPLIPRPQSGDNCSLGPQRHPINQQEMC----PRVP----SS 1727

Qy         739 PATVNNSSDTESIPSPHTEAAKDTGQNGPKPPATLGADGP-----PPGPP---------- 783
               | : :||       ||||           : |||      |     |   |
Db        1728 PQSHSNSQ------SPHTPGGHSNDGYSAQSPATPRFQSPEHCSQPSSRPHSRDAFTAVQ 1781

Qy         784 TPPRRTSRAPIEPT--PASEATGAPTPPPAPPSPSAPP--PVVPKEEKEEETAAAPPVEE 839
               | |  | ||  |:   :   | :    |    |||||    :|     : |     : :
Db        1782 KPVRSPSVAPEAPSFKNSPHHTNSTLGDPLSGKPSAPPHFSSIPSTGGFQITQQQNQMVQ 1841

Qy         840 GEEQKPPAAEELAVDTGKAEEPVKSECTEEAE------------EGPAKGKDAEAAEAT 886
               |: |:  | : :  |   |  |  |  |:|                 |  : |   |
Db        1842 GQLQQSQAQQNIGPDNYGARVPTPSG-TQEVPVVRQPDPTHQPTLPGTQEMSDISTVQDP 1900

Qy         887 AEGALK-----------------------AEKKEGGSGRATTAKSSGAPQDSDSSATC 921
               | | |                      ::||  :  |  | :: |    ::|:
Db        1901 ALGGLSPSELEKHRQRLREFLIRQQMQRNSIKQEKEAAAAAAAAAAAAAA---GNASSGW 1957

Qy         922 SADEVDEAEGGDKNRLLSPRPS--LLTPTGDPRANASPQKPLDLKQLKQRAAAIPPIQVT 979
               :  |:  :   :|  | |  :     :|  : :  |: :  |:  :  ||:
Db        1958 TGGEICAFQQDKTHRAPPPYPQDRVTMSAAGTQAPVAGKMPVAVGGLEDKLIRPPPMGTP 2017

Qy         980 KVHEPPREDAAPTKPAPP-----APPPPQNLQPESDAPQ---QPGSSPRGKSRSPAPPAD 1031
               : :|      |  |:  |     || | : |  |:   |    ||      |  :  |
Db        2018 AIMDP--NTLRPQGPSRPQGMFNRPPFPPHWQDQSTGPRRFPQPDLQAMGIRHNLNP--- 2072

Qy        1032 KEAFAAEAQKLPGDPPCWTSGLPFPV---PPREVIKASPHAPDPSAFSYAPPGHPLPL-- 1086
                  ||  | : |       | |  :      ||::     | |           |||
Db        2073 ----AANVQNMEG------LGNPHTIIAGHGGEVMQPMSQGPPPQFIELRHNAQRLPLRP 2122

Qy        1087 -----GLHDTARPVLPR----PPTIS-NPPPLISSAKHPSVLERQIGAISQGMSVQLHVP 1136
                     |    ||  :|:     | ||| :|     | :        |:|    |:|| |
Db        2123 QFMPRGPQPRARLFVPQQTMSAPYISQHPISQTGSIQTDGATNSQMGLQQGGLSVLL--- 2179

Qy        1137 YSEHAKAPVGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPESLGVPTAQEASV-- 1194
                    : | | ||    :       :|  |  | || |:     |:  |  ::|
Db        2180 ----PQQPTGSVTHKSHMGPQAASSSPNVGTVQSQLPPQSVVTRPQPTTVENSEELPEPD 2235

Qy        1195 LRGTALGSVPGGSITKGIPSTRVPSDSAITYRGSI----THGTPADVLYKGTITRIIGED 1250
               | |    |  ||  :  :  :  |      |:     |:    | |       ::
Db        2236 LEGLGDASADGGVEDEDDLALDLDPDKGDDDLGNLDNLETNDPHLDDLLNSDEFDLLAYT 2295

Qy        1251 SPSRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPHETAAPKRT 1310
               |  ||:|    ||              :  : : :  |       |    :|  |
Db        2296 DP-ELDQGD----PKD------------VFSDQLRLVEAETE------APSSGSAGVKVE 2332

Qy        1311 YDMMEGRVGRAISSASIEGLMGRAIPPERHSPHHLKEQHHIRGSITQGIPRSYVEAQEDY 1370
                  : :|:   |: | :  |:    :|    :  :|| :    | | |  ||  |  :::
Db        2333 IKVEQGQKCSAVHSTA--GVCANQLPASSKTAGNLKIKVEDGGLIPQVQPRQIV--KDEI 2388

Qy        1371 LRREAKLLKREGTPPPPPSRDLTEAYKTQALG------PLKLKPAHEGLVATVKEAGRS 1424
                  : ||    : |      :  :  :  ||      | :  | |     |  :|
Db        2389 GEAVSALLGGTTSSPKSTQPENQPASLSSVRLGGLSYPLPAQTDPLH--FPPTGSDADDD 2446

Qy        1425 IHEIPREELRHTP-------------ELPLAPRPLKE------------GSITQGTPLKY 1459
                  |:|    :|:|             ||||  : | |            |: ||    :
Db        2447 ALELPDVGGQHSPAVDLAKVESSLDGELPLLIQDLLEHEKKEQQKQQQLSSLHQGGVAPH 2506

Qy        1460 DTGASTTGSKKHDV-------------RSLIGSPGRTFPPVHPLDVMADARALERACYE 1505
                  :  ||          |          : ::| ||    | |    |: : :  ::
Db        2507 FSALSTNQQPNPQVAGQIMLPPHHRPPPQGMMGPPGMVPRPSH---VLQNQQPQQQRLMG 2563

Qy        1506 ESLKSRPGTASSSGGSIAR-GAPVIVPELGKPRQSPLTYEDHGAPFAGHLPRGSPVTMRE 1564
                  |  | | :  :: | | | |   || :|          | :|        :::
Db        2564 PGLVPPPHMAMNQQQTMIRMGQPGIHAGLGHQQQ----------PQSG---------VKQ 2604
```

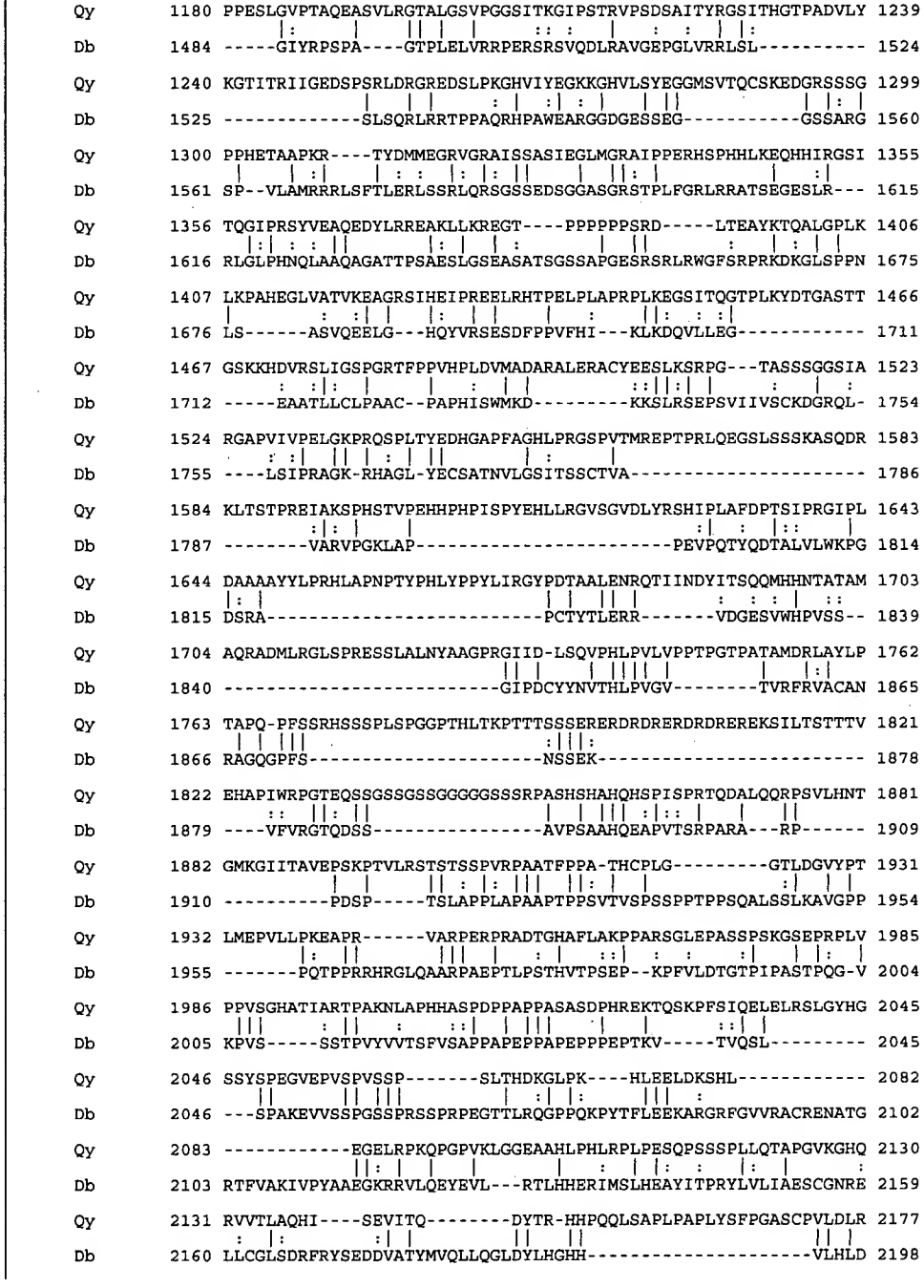
```
Qy       1565 PTPRLQEGSLSSSKASQDRKLTS-TPREIAKSPHSTVPEHHPHPISPYEHL-LRGVSGVD 1622
Db       2605 PP--------LSNNFFPDKDLDKFTTDDIM-----------DPIAKAKMVALKGINRV- 2643

Qy       1623 LYRSHIPLAFDPTSIPRGIPLDAAAAYYLPRHLA--------------------PNPTYP 1662
Db       2644 -------LAQDPMVVPPGINREQVS--LLAQRLASAPATDAGQLPSGPPKEGETSDPTQS 2694

Qy       1663 HLYPPYLIRGYPDTAALENRQTIIND---------YITSQQM--------------HHNT 1699
Db       2695 RPNPPQFVQG-----------IINDAEKHQYEEWLLHTQQLLQMQLKFLEEQIGVHRKS 2742

Qy       1700 ATAMA--------------------------------QRADMLRGLSPRESSLALNYA-- 1725
Db       2743 RKALCAKQRTAKKAGREFAEADAEKLKLVTEEQSKIQKQLDQVRKQQKEHTNLVAEYRSK 2802

Qy       1726 ---------------AGPRGIIDLSQVPH-----------------------------L 1740
Db       2803 QQQHQQSSSLLNPGHSGPAG------APHMFPKMPGQMVIGQQGAQVMGQHPTMMPQAGM 2856

Qy       1741 PVLVP--------PTPGTPATAMDRLAYLPTAPQPFSSRHSSSPLSPGGPTHLTKPTTTS 1792
Db       2857 PVRMPQGQPFVGGPQPQLPATLGNSGV---RGPGPAATPAGFLPQGPG----MQSPDARL 2909

Qy       1793 SSERERDRDRERDRDRERERKSILTSTTTVEHA------------------PIWRPGTEQS 1834
Db       2910 LQERQLQHRMQMAKLQQQQQQIMMGQQPIPHAGNSQTNLIPQTQSGMIGNPVMAQQVNAQ 2969

Qy       1835 SGSSGSSGGGGGS--------------------SSRPASHSHAHQHSP-ISPRTQDAL-- 1871
Db       2970 QGMPSNQGSTQGMMQIPQGVVGSQTVVSLPQNLAGQPIHHAQAIAGQPGIMGNQQVAMSE 3029

Qy       1872 QQRPSVL----------HNTGMKGIITAVEPSKPTVLRSTSTSSPVRPAATFPPATH--- 1918
Db       3030 QQRPMQMLSQQGMVGSPGHPGIRGPHSHLTPQQQNILAQRMLASQQQQQQQQQQQLHQQQ 3089

Qy       1919 -------------------------------CPLGGTLDGVYPTLMEPVLLPKEAPRVAR 1947
Db       3090 QQQQLHQQQQQQQLHQQQQQQQLQLQQQQQQLQQQQNVDKNMIQFQQQQQMAQKQQAMQI 3149

Qy       1948 PERPRADTGHAFLAKPPARSGLEPASSPSKGSEP--RPLVPPVSGHATIAR--------- 1996
Db       3150 SSQPSQDQG-----------GLSQPSTPQMGSSPCTRSVTPQPHGGTDSQHPCPKESGLL 3198

Qy       1997 TPAKNLAPHHASPDPPAPPASASDPHREKTQSKPFSIQELELRSLGYHGSSYSPEGVEPV 2056
Db       3199 SPESKTPPQHSGPSTPSHVYQVGSANQLQQKKDHLNLQ----KQTGLMGNQQSMVQQQQQ 3254

Qy       2057 SPVSSP----SLTHDKGLPKHLEELDKSHLEGELRPKQPGPVKLGGEAAHLPHLRPLPES 2112
Db       3255 QPLLTPQRQGSVTDDKPSMMNIKEEGKT-IDISVQQQQQQAVQ-----------NPMMQS 3302

Qy       2113 QPSSSPLL--------QTAPGVKGH--QRVVTLAQHISEVITQDYTRHHPQQLSAPLPAP 2162
Db       3303 QDSSMQLQVTGQPHPGQQQPVVMGHNPQQQALMAQHQKQQAMMGIIRAQQQGITAQRPA- 3361

Qy       2163 LYSFPGASCPVLDLRRP----------PSDLYLPPPDHGAPARGSPHSEG--GKRSPEPN 2210
Db       3362 --LQPG------QIRTPVNIQAIIAQNPQLRNLPPNQQ------IQHIQAIIAQRQIQQG 3407

Qy       2211 KTSVLGGGEDGIEPVSPPEGMTEPGHSRSAVYPLLYRDGEQTEPSRMGSKSPGNTSQPPA 2270
Db       3408 QMLRMAMGQGQIRPQMPPGQVLQVG--------------QQHQSNM--LQPGVNSQ--- 3447

Qy       2271 FFSKLTESNSAMVKSKKQEINKKLNTHNRNEPEYNISQPGTEIFNMPAITGTGLMTYRSQ 2330
Db       3448 -------MQQGMVVHGQQQ-----QSHT-GEMMQNIS-------------------RSQ 3474

Qy       2331 AVQEHASTNMGLEAIIRKALMGKYDQWEESPP--LSANAFNPLNASASLPAAMPITAADG 2388
Db       3475 APVPPATAEQG-----RMAMPASPCQPLANPPGDPQRHAFNQ-NMAMRPP-------TPN 3521

Qy       2389 RSDHTLTSPGGGGKAKVSGRPSSRKAKSP-APGLASGDRPPSVSSVHS------EGDCNR 2441
Db       3522 QNQQALMAAGG----RVQGSPS--HAYSPRGPFGMSPVHPASPNSSHASSPSMGDGRAGR 3575

Qy       2442 RTPLTNRVWEDRPSSAGSTPFPYNPLIMRLQ---AGVMASPPPPGLPAGS---------- 2488
Db       3576 GSPY-NQI-KASPLRSPGAKSPLDSLVLKVETQTSGNETSQTALGIPNGPQKSINIKQQT 3633

Qy       2489 ---GPLAGPH 2495
Db       3634 QQVSEVLGPH 3643
```

RESULT 78
US-10-333-314-18
; Sequence 18, Application US/10333314
; Publication No. US20030211093A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; YUE, Henry
; APPLICANT: KHAN, Farrah A.; GURURAJAN, Rajagopal
; APPLICANT: HAFALIA, April J.A.; CHAWLA, Narinder K.
; APPLICANT: ARVIZU, Chandra S.; RAMKUMAR, Jayalaxmi
; APPLICANT: GANDHI, Ameena R.; POLICKY, Jennifer L.
; APPLICANT: BAUGHN, Mariah R.; TRIBOULEY, Catherine M.
; APPLICANT: THORNTON, Michael B.; BANDMAN, Olga
; APPLICANT: NGUYEN, Danniel B.; LU, Yan
; APPLICANT: BURFORD, Neil; LAL, Preeti G.
; APPLICANT: DING, Li; YAO, Monique G.
; APPLICANT: ELLIOTT, Vicki S.; RECIPON, Shirley A.
; APPLICANT: KEARNEY, Liam; LU, Dyung Aina M.
; APPLICANT: GREENWALD, Sara R.; TANG, Y. Tom
; APPLICANT: XU, Yuming; WALSH, Roderick T.
; APPLICANT: GIETZEN, Kimberly J.; YANG, Junming
; APPLICANT: JACKSON, Jennifer L.
; TITLE OF INVENTION: HUMAN KINASES
; FILE REFERENCE: PI-0162 USN
; CURRENT APPLICATION NUMBER: US/10/333,314
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: PCT/US01/23092
; PRIOR FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/220,038
; PRIOR FILING DATE: 2000-07-21
; PRIOR APPLICATION NUMBER: US 60/222,112
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/222,831
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/224,729
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PERL Program
; SEQ ID NO 18
; LENGTH: 2380
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 7477141CD1
US-10-333-314-18

Query Match 2.9%; Score 377; DB 12; Length 2380;
Best Local Similarity 19.2%; Pred. No. 1e-08;
Matches 484; Conservative 295; Mismatches 886; Indels 850; Gaps 117;

```
Qy         78 RSQELHLRPESHSYLPE---LGKSEMEFIESKRPRLELLPDPLLRPSPLLATGQPAGSED 134
Db        420 RSLDMAIDPDSLTYTVQHQVLGSDQWTALVTG----------LREPGWAATGLRKGVQH 468

Qy        135 LTKDRSLTGKLEPVSPPSPP-------HTDPELELVPPRLSKEELIQNMDRVDREITMVE 187
Db        469 IFRVLSTTVKSS--SKPSPPSEPVQLLEHGPTLEEAPAMLDKPDIVY----------VVE 516

Qy        188 QQISKLKKKQQQLEEEAAKPPEPEKPVSPPPIESKHRSLVQIIYDENRKKAEAAHRILEG 247
```



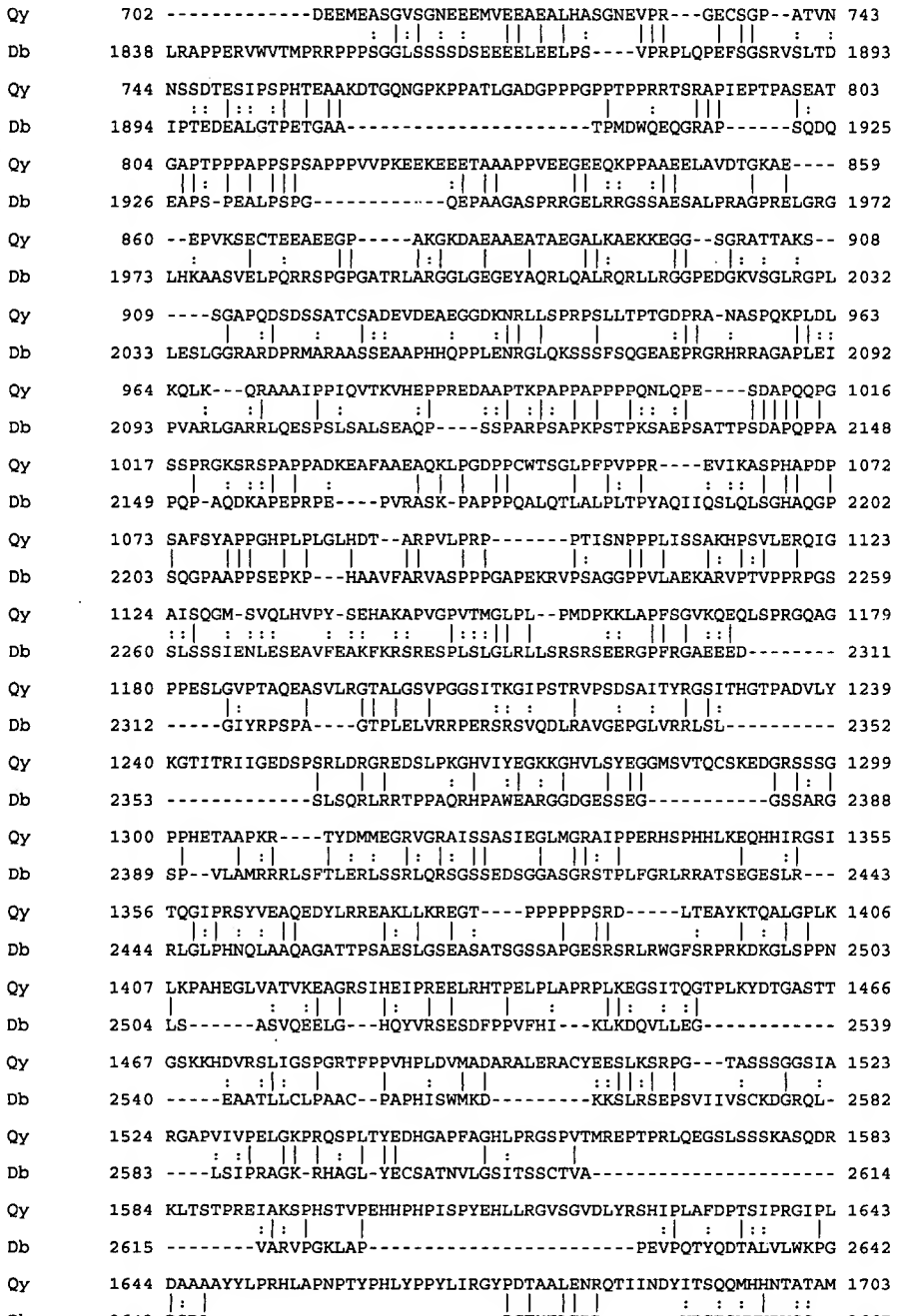


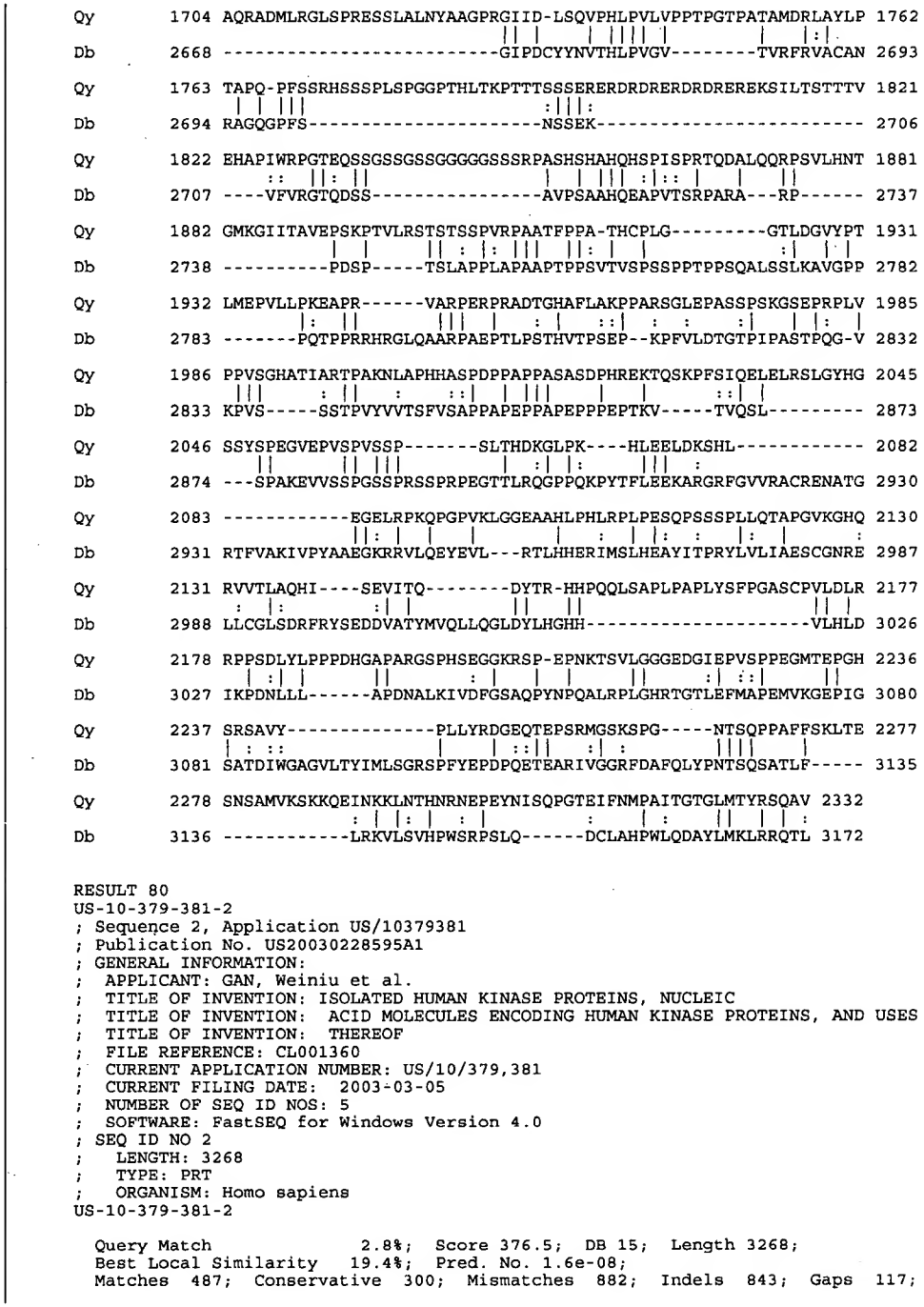
```
Qy       2178 RPPSDLYLPPPDHGAPARGSPHSEGGKRSP-EPNKTSVLGGGEDGIEPVSPPEGMTEPGH 2236
Db       2199 IKPDNLLL------APDNALKIVDFGSAQPYNPQALRPLGHRTGTLEFMAPEMVKGEPIG 2252

Qy       2237 SRSAVY--------------PLLYRDGEQTEPSRMGSKSPG-----NTSQPPAFFSKLTE 2277
Db       2253 SATDIWGAGVLTYIMLSGRSPFYEPDPQETEARIVGGRFDAFQLYPNTSQSATLF----- 2307

Qy       2278 SNSAMVKSKKQEINKKLNTHNRNEPEYNISQPGTEIFNMPAITGTGLMTYRSQAV 2332
Db       2308 -----------LRKVLSVHPWSRPSLQ------DCLAHPWLQDAYLMKLRRQTL 2344
```

RESULT 79
US-10-210-130-38
; Sequence 38, Application US/10210130
; Publication No. US20040014053A1
; GENERAL INFORMATION:
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Patturajan, Meera
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Miller, Charles E.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Li, Li
; APPLICANT: Berghs, Constance
; APPLICANT: Zhong, Mei
; APPLICANT: Casman, Stacie J.
; APPLICANT: Voss, Edward Z.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Smithson, Glennda
; APPLICANT: Ji, Weizhen
; APPLICANT: Gorman, Linda
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Leite, Mario W.
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Anderson, David W.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Khramtsov, Nikolai V.
; APPLICANT: Ort, Tatiana
; APPLICANT: Ellerman, Karen
; APPLICANT: Rastelli, Luca
; APPLICANT: Agee, Michele L.
; APPLICANT: Chaudhuri, Amitabha
; APPLICANT: Chant, John S.
; APPLICANT: DiPippo, Vincent A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Eisen, Andrew J.
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Giot, Loic
; APPLICANT: Ooi, Chean Eng
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Spaderna, Steven K.
; APPLICANT: Hjalt, Tord
; APPLICANT: Liu, Xiaohong
; APPLICANT: Taupier, Raymond J., Jr.
; APPLICANT: Catterton, Elina
; APPLICANT: Shenoy, Suresh G.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-416C (Cura-716 SMT)
; CURRENT APPLICATION NUMBER: US/10/210,130
; CURRENT FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: 60/309,501
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 60/316,508
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: 60/354,655
; PRIOR FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: 60/310,291
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 60/383,887
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: 60/310,951
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/323,936
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: 60/381,039
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/311,292
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/311,979
; PRIOR FILING DATE: 2001-08-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 38
; LENGTH: 3208
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-210-130-38

Query Match 2.9%; Score 377; DB 15; Length 3208;
Best Local Similarity 19.2%; Pred. No. 1.5e-08;
Matches 484; Conservative 295; Mismatches 886; Indels 850; Gaps 117;

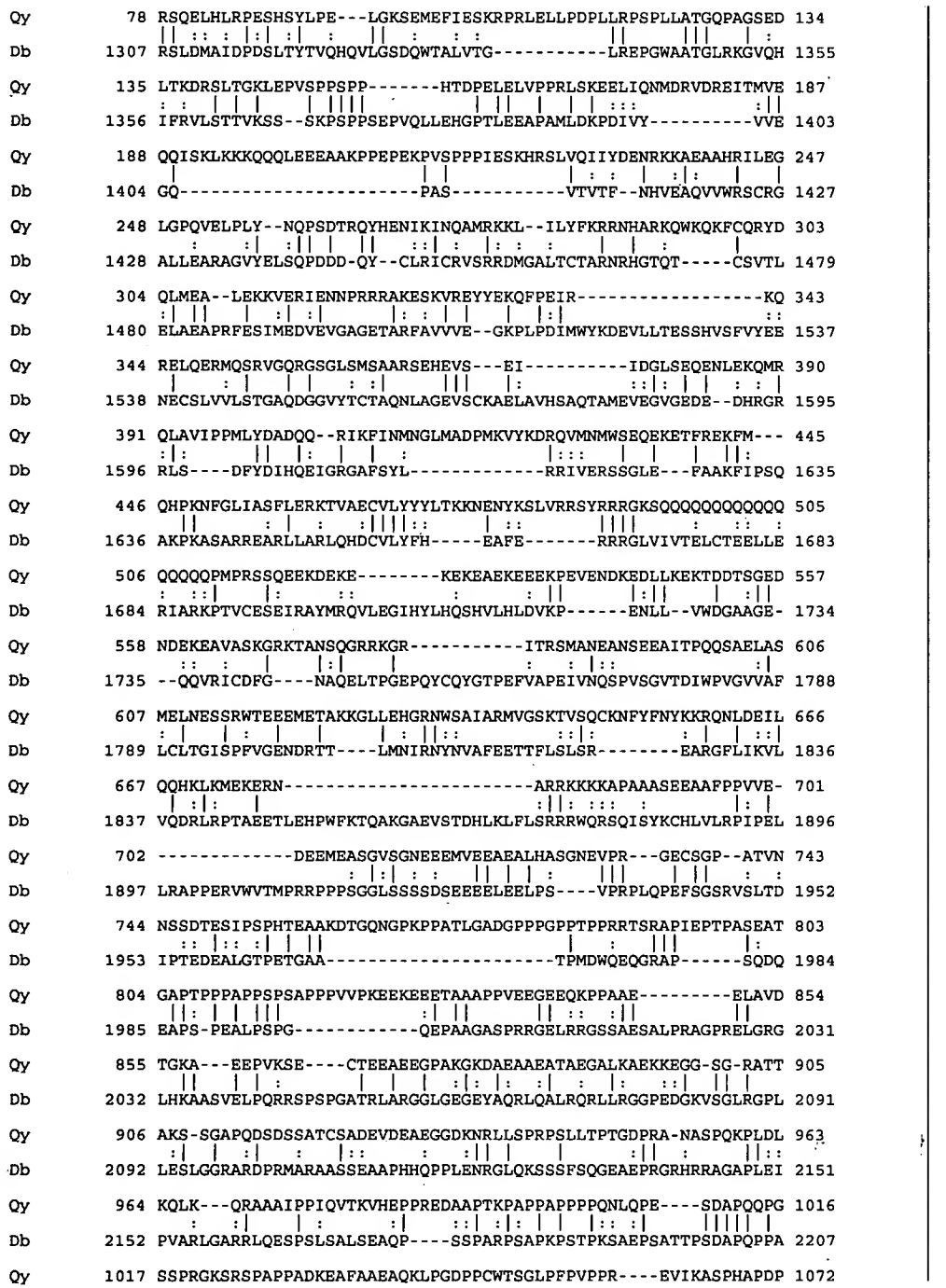
```
Qy         78 RSQELHLRPESHSYLPE---LGKSEMEFIESKRPRLELLPDPLLRPSPLLATGQPAGSED 134
Db       1248 RSLDMAIDPDSLTYTVQHQVLGSDQWTALVTG-----------LREPGWAATGLRKGVQH 1296

Qy        135 LTKDRSLTGKLEPVSPPSPP-------HTDPELELVPPRLSKEELIQNMDRVDREITMVE 187
Db       1297 IFRVLSTTVKSS--SKPSPPSEPVQLLEHGPTLEEAPAMLDKPDIVY----------VVE 1344

Qy        188 QQISKLKKKQQQLEEEAAKPPEPEKPVSPPPIESKHRSLVQIIYDENRKKAEAAHRILEG 247
Db       1345 GQ----------------------PAS----------VTVTF--NHVEAQVVWRSCRG 1368

Qy        248 LGPQVELPLY--NQPSDTRQYHENIKINQAMRKKL--ILYFKRRNHARKQWKQKFCQRYD 303
Db       1369 ALLEARAGVYELSQPDDD-QY--CLRICRVSRRDMGALTCTARNRHGTQT-----CSVTL 1420

Qy        304 QLMEA--LEKKVERIENNPRRRAKESKVREYYEKQFPEIR----------------KQ 343
Db       1421 ELAEAPRFESIMEDVEVGAGETARFAVVVE--GKPLPDIMWYKDEVLLTESSHVSFVYEE 1478

Qy        344 RELQERMQSRVGQRGSGLSMSAARSEHEVS---EI----------IDGLSEQENLEKQMR 390
Db       1479 NECSLVVLSTGAQDGGVYTCTAQNLAGEVSCKAELAVHSAQTAMEVEGVGEDE--DHRGR 1536

Qy        391 QLAVIPPMLYDADQQ--RIKFINMNGLMADPMKVYKDRQVMNMWSEQEKETFREKFM--- 445
Db       1537 RLS----DFYDIHQEIGRGAFSYL------------RRIVERSSGLE---FAAKFIPSQ 1576

Qy        446 QHPKNFGLIASFLERKTVAECVLYYYLTKKNENYKSLVRRSYRRRGKSQQQQQQQQQQQQ 505
Db       1577 AKPKASARREARLLARLQHDCVLYFH-----EAFE-------RRRGLVIVTELCTEELLE 1624

Qy        506 QQQQQPMPRSSQEEKDEKE--------KEKEAEKEEEKPEVENDKEDLLKEKTDDTSGED 557
Db       1625 RIARKPTVCESEIRAYMRQVLEGIHYLHQSHVLHLDVKP------ENLL--VWDGAAGE- 1675

Qy        558 NDEKEAVASKGRKTANSQGRRKGR----------ITRSMANEANSEEAITPQQSAELAS 606
Db       1676 --QQVRICDFG----NAQELTPGEPQYCQYGTPEFVAPEIVNQSPVSGVTDIWPVGVVAF 1729

Qy        607 MELNESSRWTEEEMETAKKGLLEHGRNWSAIARMVGSKTVSQCKNFYFNYKKRQNLDEIL 666
Db       1730 LCLTGISPFVGENDRTT----LMNIRNYNVAFEETTFLSLSR--------EARGFLIKVL 1777

Qy        667 QQHKLKMEKERN----------------------ARRKKKKAPAAASEEAAFPPVVE- 701
Db       1778 VQDRLRPTAEETLEHPWFKTQAKGAEVSTDHLKLFLSRRRWQRSQISYKCHLVLRPIPEL 1837
```





RESULT 80
US-10-379-381-2
; Sequence 2, Application US/10379381
; Publication No. US20030228595A1
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001360
; CURRENT APPLICATION NUMBER: US/10/379,381
; CURRENT FILING DATE: 2003-03-05
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 3268
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-379-381-2

Query Match 2.8%; Score 376.5; DB 15; Length 3268;
Best Local Similarity 19.4%; Pred. No. 1.6e-08;
Matches 487; Conservative 300; Mismatches 882; Indels 843; Gaps 117;


Qy 78 RSQELHLRPESHSYLPE---LGKSEMEFIESKRPRLELLPDPLLRPSPLLATGQPAGSED 134
Db 1307 RSLDMAIDPDSLTYTVQHQVLGSDQWTALVTG----------LREPGWAATGLRKGVQH 1355
Qy 135 LTKDRSLTGKLEPVSPPSPP-------HTDPELELVPPRLSKEELIQNMDRVDREITMVE 187
Db 1356 IFRVLSTTVKSS--SKPSPPSEPVQLLEHGPTLEEAPAMLDKPDIVY----------VVE 1403
Qy 188 QQISKLKKKQQQLEEEAAKPPEPEKPVSPPPIESKHRSLVQIIYDENRKKAEAAHRILEG 247
Db 1404 GQ----------PAS----------VTVTF--NHVEAQVVWRSCRG 1427
Qy 248 LGPQVELPLY--NQPSDTRQYHENIKINQAMRKKL--ILYFKRRNHARKQWKQKFCQRYD 303
Db 1428 ALLEARAGVYELSQPDDD-QY--CLRICRVSRRDMGALTCTARNRHGTQT-----CSVTL 1479
Qy 304 QLMEA--LEKKVERIENNPRRRAKESKVREYYEKQFPEIR----------KQ 343
Db 1480 ELAEAPRFESIMEDVEVGAGETARFAVVVE--GKPLPDIMWYKDEVLLTESSHVSFVYEE 1537
Qy 344 RELQERMQSRVGQRGSGLSMSAARSEHEVS---EI----------IDGLSEQENLEKQMR 390
Db 1538 NECSLVVLSTGAQDGGVYTCTAQNLAGEVSCKAELAVHSAQTAMEVEGVGEDE--DHRGR 1595
Qy 391 QLAVIPPMLYDADQQ--RIKFINMNGLMADPMKVYKDRQVMNMWSEQEKETFREKFM--- 445
Db 1596 RLS----DFYDIHQEIGRGAFSYL----------RRIVERSSGLE---FAAKFIPSQ 1635
Qy 446 QHPKNFGLIASFLERKTVAECVLYYYLTKKNENYKSLVRRSYRRRGKSQQQQQQQQQQQQ 505
Db 1636 AKPKASARREARLLARLQHDCVLYFH-----EAFE-------RRRGLVIVTELCTEELLE 1683
Qy 506 QQQQQPMPRSSQEEKDEKE--------KEKEAEKEEEKPEVENDKEDLLKEKTDDTSGED 557
Db 1684 RIARKPTVCESEIRAYMRQVLEGIHYLHQSHVLHLDVKP------ENLL--VWDGAAGE- 1734
Qy 558 NDEKEAVASKGRKTANSQGRRKGR----------ITRSMANEANSEEAITPQQSAELAS 606
Db 1735 --QQVRICDFG----NAQELTPGEPQYCQYGTPEFVAPEIVNQSPVSGVTDIWPVGVVAF 1788
Qy 607 MELNESSRWTEEEMETAKKGLLEHGRNWSAIARMVGSKTVSQCKNFYFNYKKRQNLDEIL 666
Db 1789 LCLTGISPFVGENDRTT----LMNIRNYNVAFEETTFLSLSR--------EARGFLIKVL 1836
Qy 667 QQHKLKMEKERN----------ARRKKKKAPAAASEEAAFPPVVE- 701
Db 1837 VQDRLRPTAEETLEHPWFKTQAKGAEVSTDHLKLFLSRRRWQRSQISYKCHLVLRPIPEL 1896
Qy 702 ----------DEEMEASGVSGNEEEMVEEAEALHASGNEVPR---GECSGP--ATVN 743
Db 1897 LRAPPERVWVTMPRRPPPSGGLSSSSDSEEEELEELPS----VPRPLQPEFSGSRVSLTD 1952
Qy 744 NSSDTESIPSPHTEAAKDTGQNGPKPPATLGADGPPPGPPTPPRRTSRAPIEPTPASEAT 803
Db 1953 IPTEDEALGTPETGAA----------TPMDWQEQGRAP------SQDQ 1984
Qy 804 GAPTPPPAPPSPSAPPPVVPKEEKEEETAAAPPVEEGEEQKPPAAE---------ELAVD 854
Db 1985 EAPS-PEALPSPG----------QEPAAGASPRRGELRRGSSAESALPRAGPRELGRG 2031
Qy 855 TGKA---EEPVKSE----CTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGG-SG-RATT 905
Db 2032 LHKAASVELPQRRSPSPGATRLARGGLGEGEYAQRLQALRQRLLRGGPEDGKVSGLRGPL 2091
Qy 906 AKS-SGAPQDSDSSATCSADEVDEAEGGDKNRLLSPRPSLLTPTGDPRA-NASPQKPLDL 963
Db 2092 LESLGGRARDPRMARAASSEAAPHHQPPLENRGLQKSSSFSQGEAEPRGRHRRAGAPLEI 2151
Qy 964 KQLK---QRAAAIPPIQVTKVHEPPREDAAPTKPAPPAPPPPQNLQPE----SDAPQQPG 1016
Db 2152 PVARLGARRLQESPSLSALSEAQP----SSPARPSAPKPSTPKSAEPSATTPSDAPQPPA 2207
Qy 1017 SSPRGKSRSPAPPADKEAFAAEAQKLPGDPPCWTSGLPFPVPPR----EVIKASPHAPDP 1072
Db 2208 PQP-AQDKAPEPRPE----PVRASK-PAPPPQALQTLALPLTPYAQIIQSLQLSGHAQGP 2261
Qy 1073 SAFSYAPPGHPLPLGLHDT--ARPVLPRP-------PTISNPPPLISSAKHPSVLERQIG 1123
Db 2262 SQGPAAPPSEPKP---HAAVFARVASPPPGAPEKRVPSAGGPPVLAEKARVPTVPPRPGS 2318
Qy 1124 AISQGM-SVQLHVPY-SEHAKAPVGPVTMGLPL--PMDPKKLAPFSGVKQEQLSPRGQAG 1179
Db 2319 SLSSSIENLESEAVFEAKFKRSRESPLSLGLRLLSRSRSEERGPFRGAEEED-------- 2370
Qy 1180 PPESLGVPTAQEASVLRGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHGTPADVLY 1239
Db 2371 -----GIYRPSPA----GTPLELVRRPERSRSVQDLRAVGEPGLVRRLSL---------- 2411
Qy 1240 KGTITRIIGEDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSSG 1299
Db 2412 ------------SLSQRLRRTPPAQRHPAWEARGGDGESSEG----------GSSARG 2447
Qy 1300 PPHETAAPKR----TYDMMEGRVGRAISSASIEGLMGRAIPPERHSPHHLKEQHHIRGSI 1355
Db 2448 SP--VLAMRRRLSFTLERLSSRLQRSGSSEDSGGASGRSTPLFGRLRRATSEGESLR--- 2502
Qy 1356 TQGIPRSYVEAQEDYLRREAKLLKREGT----PPPPPPSRD-----LTEAYKTQALGPLK 1406
Db 2503 RLGLPHNQLAAQAGATTPSAESLGSEASATSGSSAPGESRSRLRWGFSRPRKDKGLSPPN 2562
Qy 1407 LKPAHEGLVATVKEAGRSIHEIPREELRHTPELPLAPRPLKEGSITQGTPLKYDTGASTT 1466
Db 2563 LS------ASVQEELG---HQYVRSESDFPPVFHI---KLKDQVLLEG---------- 2598
Qy 1467 GSKKHDVRSLIGSPGRTFPPVHPLDVMADARALERACYEESLKSRPG---TASSSGGSIA 1523
Db 2599 -----EAATLLCLPAAC--PAPHISWMKD---------KKSLRSEPSVIIVSCKDGRQL- 2641
Qy 1524 RGAPVIVPELGKPRQSPLTYEDHGAPFAGHLPRGSPVTMREPTPRLQEGSLSSSKASQDR 1583
Db 2642 ----LSIPRAGK-RHAGL-YECSATNVLGSITSSCTVA---------- 2673
Qy 1584 KLTSTPREIAKSPHSTVPEHHPHPISPYEHLLRGVSGVDLYRSHIPLAFDPTSIPRGIPL 1643
Db 2674 --------VARVPGKLAP----------PEVPQTYQDTALVLWKPG 2701
Qy 1644 DAAAAYYLPRHLAPNPTYPHLYPPYLIRGYPDTAALENRQTIINDYITSQQMHHNTATAM 1703
Db 2702 DSRA----------PCTYTLERR-------VDGESVWHPVSS-- 2726
Qy 1704 AQRADMLRGLSPRESSLALNYAAGPRGIID-LSQVPHLPVLVPPTPGTPATAMDRLAYLP 1762
Db 2727 ----------GIPDCYYNVTHLPVGV--------TVRFRVACAN 2752
Qy 1763 TAPQ-PFSSRHSSSPLSPGGPTHLTKPTTTSSSERERDRDRERDRDRDREREKSILTSTTTV 1821
Db 2753 RAGQGPFS----------NSSEK---------- 2765
Qy 1822 EHAPIWRPGTEQSSGSSGSSGGGGSSSRPASHSHAHQHSPISPRTQDALQQRPSVLHNT 1881
Db 2766 ----VFVRGTQDSS----------AVPSAAHQEAPVTSRPARA---RP------ 2796
Qy 1882 GMKGIITAVEPSKPTVLRSTSTSSPVRPAATFPPA-THCPLG---------GTLDGVYPT 1931
Db 2797 ----------PDSP-----TSLAPPLAPAAPTPPSVTVSPSSPPTPPSQALSSLKAVGPP 2841
Qy 1932 LMEPVLLPKEAPR------VARPERPRADTGHAFLAKPPARSGLEPASSPSKGSEPRPLV 1985
Db 2842 -------PQTPPRRHRGLQAARPAEPTLPSTHVTPSEP--KPFVLDTGTPIPASTPQG-V 2891
Qy 1986 PPVSGHATIARTPAKNLAPHHASPDPPAPPASASDPHREKTQSKPFSIQELELRSLGYHG 2045
Db 2892 KPVS-----SSTPVYVVTSFVSAPPAPEPPAPEPPPEPTKV-----TVQSL--------- 2932
Qy 2046 SSYSPEGVEPVSPVSSP------SLTHDKGLPKHLEELDKSHLEGELRPKQPGPVKLGG 2098

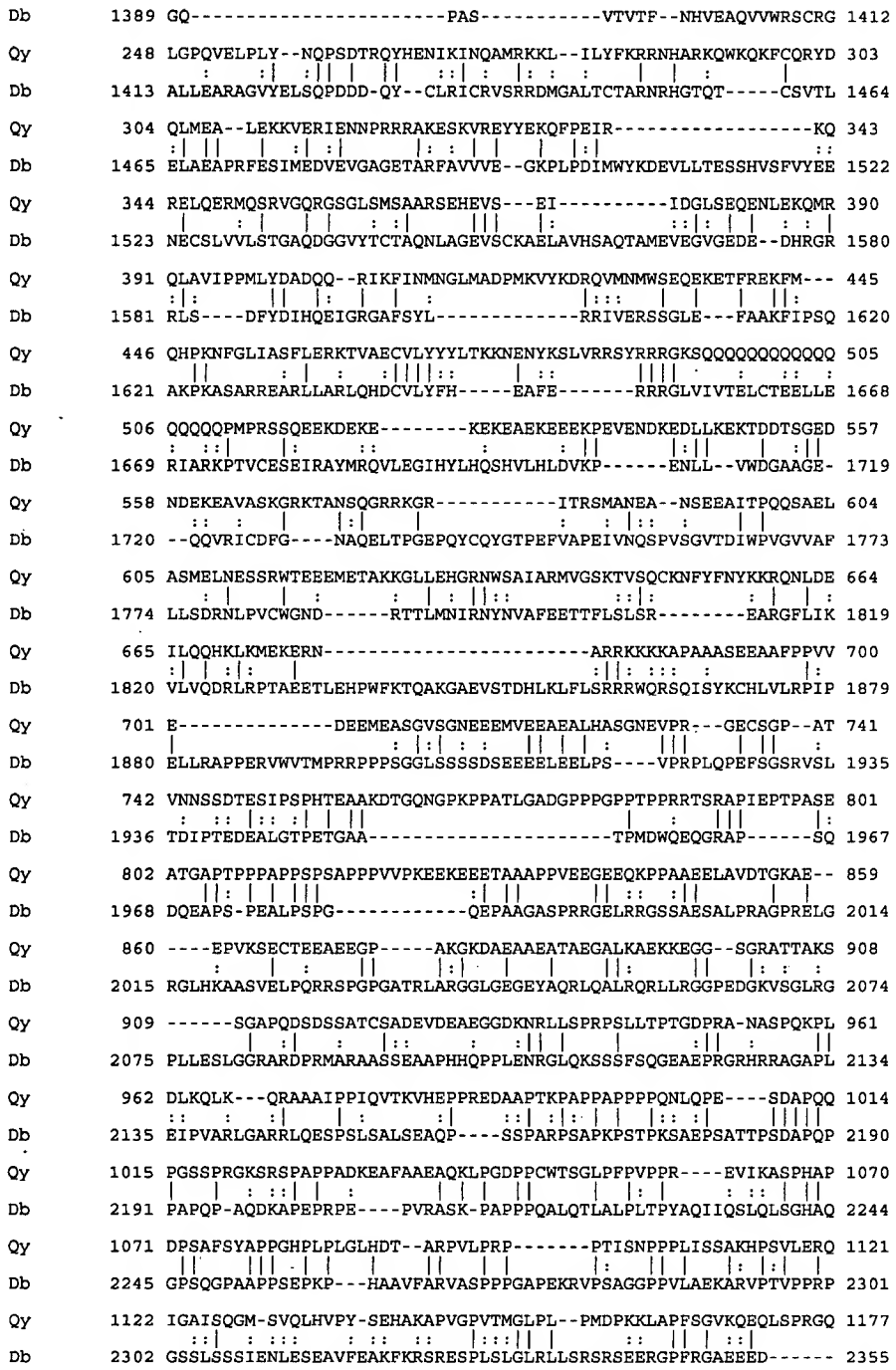
```
Db       2933 ---SPAKEVVSSPGSSPRSSPRPEGTTLRQGPP----QKPYTFLEEKARQGRFGVVRACR 2985

Qy       2099 EAAHLPHLRPLPESQPSSSPLLQTAPGVKGHQRVVTLAQHISEVITQDY----TRHHPQQ 2154

Db       2986 E---------------------NATGRTFVAKIVPYAAEGKRRVLQEYEVLRTLHHERI 3023

Qy       2155 LS---APLPAPLYSFPGASCPVLDLRRPPSDLYLPPPD----------------HG---- 2191

Db       3024 MSLHEAYITPRYLVLIAESCGNRELLCGLSDRFRYSEDDVATYMVQLLQGLDYLHGHHVL 3083

Qy       2192 -----------APARGSPHSEGGKRSP-EPNKTSVLGGGEDGIEPVSPPEGMTEPGHSRS 2239

Db       3084 HLDIKPDNLLLAPDNALKIVDFGSAQPYNPQALRPLGHRTGTLEFMAPEMVKGEPIGSAT 3143

Qy       2240 AVY--------------PLLYRDGEQTEPSRMGSKSPG-----NTSQPPAFFSKLTESNS 2280

Db       3144 DIWGAGVLTYIMLSGRSPFYEPDPQETEARIVGGRFDAFQLYPNTSQSATLF-------- 3195

Qy       2281 AMVKSKKQEINKKLNTHNRNEPEYNISQPGTEIFNMPAITGTGLMTYRSQAV 2332

Db       3196 ---------LRKVLSVHPWSRPSLQ------DCLAHPWLQDAYLMKLRRQTL 3232


RESULT 81
US-10-092-900A-224
; Sequence 224, Application US/10092900A
; Publication No. US20040043382A1
; GENERAL INFORMATION:
;  APPLICANT: Padigaru, Muralidhara
;  APPLICANT:  Spytek, Kimberly A.
;  APPLICANT:  Shenoy, Suresh G.
;  APPLICANT:  Taupier Jr., Raymond J.
;  APPLICANT:  Pena, Carol E.A.
;  APPLICANT:  Li, Li
;  APPLICANT:  Zerhusen, Bryan D.
;  APPLICANT:  Gusev, Vladimir Y.
;  APPLICANT:  Ji, Weizhen
;  APPLICANT:  Gorman, Linda
;  APPLICANT:  Miller, Charles E.
;  APPLICANT:  Kekuda, Ramesh
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Gangolli, Esha A.
;  APPLICANT:  Vernet, Corine A.M.
;  APPLICANT:  Guo, Xiaojia Sasha
;  APPLICANT:  Tchernev,, Velizar T.
;  APPLICANT:  Fernandes, Elma R.
;  APPLICANT:  Casman, Stacie J.
;  APPLICANT:  Malyankar, Uriel M.
;  APPLICANT:  Gerlach, Valerie
;  APPLICANT:  Liu, Yi
;  APPLICANT:  Anderson, David W.
;  APPLICANT:  Spaderna, Steven K.
;  APPLICANT:  Catterton, Elina
;  APPLICANT:  Leite, Mario W.
;  APPLICANT:  Zhong, Haihong
;  APPLICANT:  Alsobrook, John P.
;  APPLICANT:  Lepley, Denise M.
;  APPLICANT:  Rieger, Daniel K.
;  APPLICANT:  Burgess, Catherine E.
;  TITLE OF INVENTION: No. US20040043382A1el Proteins and Nucleic Acids Encoding Same
;  FILE REFERENCE: 21402-290C
;  CURRENT APPLICATION NUMBER: US/10/092,900A
;  CURRENT FILING DATE:  2002-03-07
;  PRIOR APPLICATION NUMBER: USSN 60/274,322
;  PRIOR FILING DATE: 2001-03-08
;  PRIOR APPLICATION NUMBER: USSN 60/283,675
;  PRIOR FILING DATE: 2001-04-13
;  PRIOR APPLICATION NUMBER: USSN 60/338,092
;  PRIOR FILING DATE: 2001-12-03
;  PRIOR APPLICATION NUMBER: USSN 60/274,281
;  PRIOR FILING DATE: 2001-03-08
;  PRIOR APPLICATION NUMBER: USSN 60/274,191
;  PRIOR FILING DATE: 2001-03-08
;  PRIOR APPLICATION NUMBER: USSN 60/325,681
;  PRIOR FILING DATE: 2001-09-27
;  PRIOR APPLICATION NUMBER: USSN 60/304,354
;  PRIOR FILING DATE: 2001-07-10
;  PRIOR APPLICATION NUMBER: USSN 60/279,995
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: USSN 60/294,899
;  PRIOR FILING DATE: 2001-05-31
;  PRIOR APPLICATION NUMBER: USSN 60/287,424
;  PRIOR FILING DATE: 2001-04-30
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 768
; SEQ ID NO 224
;   LENGTH: 1390
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-092-900A-224

  Query Match             2.8%;  Score 375.5;  DB 12;  Length 1390;
  Best Local Similarity   19.1%;  Pred. No. 6.4e-09;
  Matches  282;  Conservative  170;  Mismatches  517;  Indels  507;  Gaps   64;

Qy        104 ESKRPRLELLPDPLLRPSPLLATGQPAGSEDLTKDRSLTGKLEPVSPPSPPHTDPEL--E 161

Db        201 QASTPPQTQTPQP--NPPPVQATPHPFPA--VTPDLIVQTPVMTVVPPQPLQTPPPVPPQ 256

Qy        162 LVPPRLSKEELIQNMDRVDREITMVEQQISKLKKKQQQ----------LEEEAAKPPEP 210

Db        257 PQPPPAPAPQPVQSHPPI---IAATPQPVKTKKGVKRKADTTTPTTIDPIHEPPSLPPEP 313

Qy        211 E----------KPVSPP----PIESKHRSLVQIIYDENRKKAEAAHRILEGLGPQVELP 255

Db        314 KTTKLGQRRESSRPVKPPKKDVPDSQQHPA---------------------------- 343

Qy        256 LYNQPSDTRQYHENIKINQAMRKKLILYFKRRNHARKQWK----------QKFCQ--R 301

Db        344 ----PEKSSKVSEQLKCCSGILKEMF----AKKHAAYAWPFYKPVDVEALGLHDYCDIIK 395

Qy        302 YDQLMEALEKKVERIENNPRRRAKE---------SKVREYYEKQFPEIRKQRELQERMQS 352

Db        396 HPMDMSTIKSKLEARE---YRDAQEFGADVRLMFSNCYKYNPPDHEVVAMARKLQDVFEM 452

Qy        353 RVGQRGSGLSMSAARSEHEVSEIIDGLSEQENLEKQMRQLAVIPPMLYDADQQRIKFINM 412

Db        453 RF----------AKMPDEPEEPVVAVSS----------PAVPP--------------- 475

Qy        413 NGLMADPMKVYKDRQVMNMWSEQEKETFREKFMQHPKNFGLIASFLERKTVAECVLYYYL 472

Db        476 ------PTKVVAPPSSSDSSSDSSSDSDSS---------------------------- 499

Qy        473 TKKNENYKSLVRRSYRRRGKSQQQQQQQQQQ----QQQQQQQPMPRSSQEEKDEKEKEKE 528

Db        500 TDDSEEERA------QRLAELQEQLKAVHEQLAALSQPQQNKP----KKKEKDKKEKKK- 548

Qy        529 AEKEEEKPEVENDKEDLLKE------KTDDTSGEDNDEKEAVASKGRKTANSQGRRKGRI 582

Db        549 -EKHKRKEEVEENKKSKAKEPPPKKTKKNNSSNSNVSKKEPAPMKSKPPP---------- 597

Qy        583 TRSMANEANSEEAITPQQSAE--LASMELNESSRWTEEEMETAKKGLLEHGRNWSAIARM 640

Db        598 ----TYESEEEDKCKPMSYEEKRQLSLDINKLP----------------GEKLGRVVHI 636

Qy        641 VGSKTVSQCKNFYFNYKKRQNLDEI-LQQHKLKMEKERNARR------KKKKAPAAASEE 693

Db        637 IQSREPS---------LKNSNPDEIEIDFETLKPSTLRELERYVTSCLRKKKRKPQAEKVD 687

Qy        694 AAFPPVVEDEEMEASGVSGNEEEMVEEAEALHASGNEVPRGECSGPATVNNSSDTESIPS 753

Db        688 V----IAGSSKMK--GFSSSESESSSESSS----------------SDSEDSETEMA 722

Qy        754 PHTEAAKDTG--------------QNGPKPPATLGADGPPPGPPTPPRRTSRAPIEPTPA 799

Db        723 PKSKKKGHPGREQKQHHHHHHQQMQQAPAPV----PQQPPPPPQQPP---------PPPP 769
```

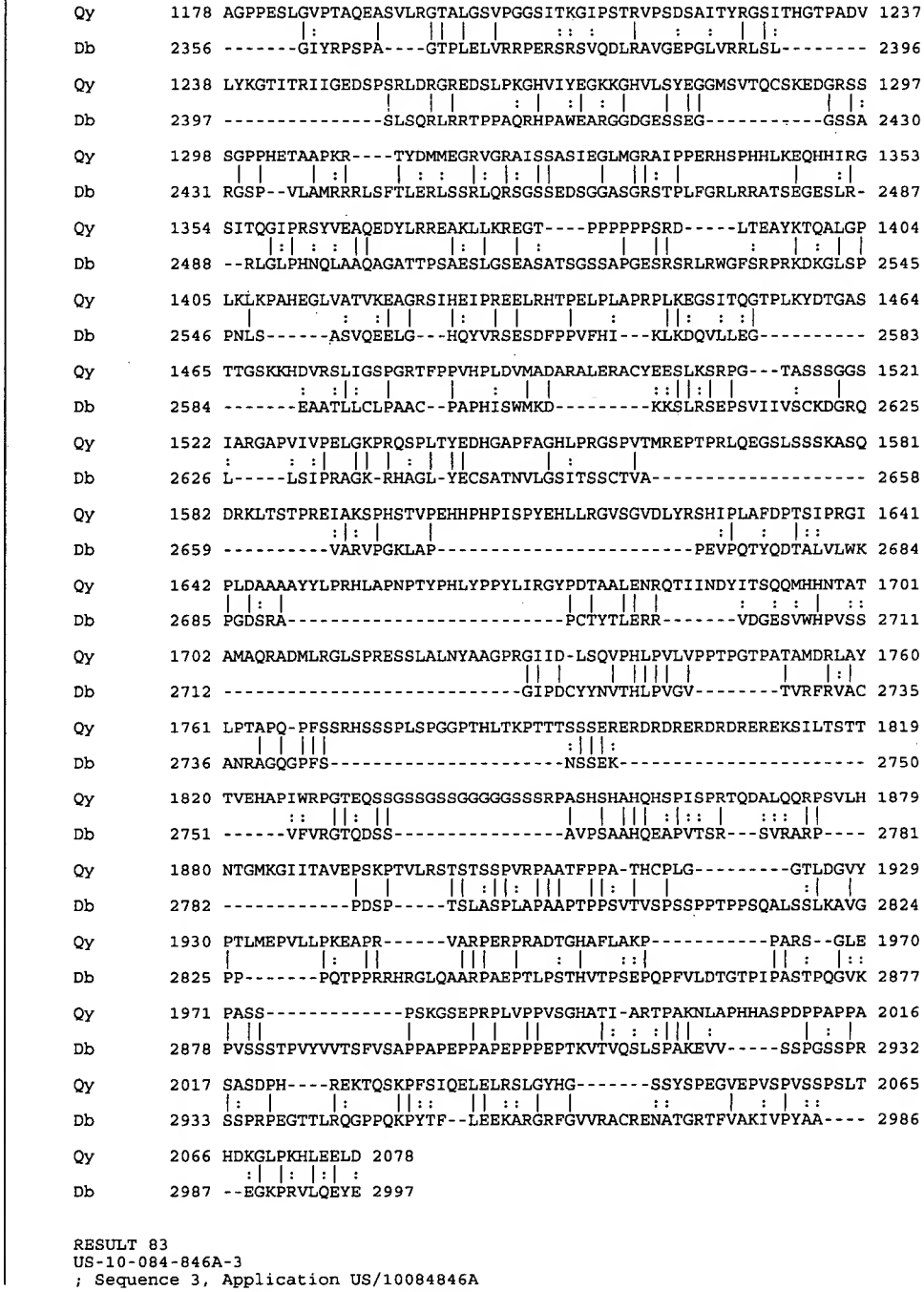
```
Qy        800 SEATGAPTPPPAPPS---------PSAPPPV----VPKEEKEEETAAAPPVEEGEEQKPP 846
Db        770 PQQQQQPPPPPPPPSMPQQAAPAMKSSPPPFIATQVPVLEPQLPGSVFDPI--GHFTQP- 826

Qy        847 AAEELAVDTGKAEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGSGRATTA 906
Db        827 -----ILHLPQPELPPHLPQPPEHSTPPHLNQHAVVSPPALHNAL-PQQPSRPSNRAAAL 880

Qy        907 KSSGAPQDSDSSATCSADEVDEAEGGDKNRLLSPRPSLLTP----TGDPRANASP----Q 958
Db        881 PPKPARPPAVSPALTQTP------------LLPQPPMAQPPQVLLEDEEPPAPPLTSMQ 927

Qy        959 KPLDLKQLKQRAAAIPPIQVTKVHEPPREDAAPTKPAPPAPPPP-----QNL--QPESDA 1011
Db        928 MQLYLQQLQKVQPPTPLLPSVKVQ---------SQPPPPLPPPPHPSVQQQLQQQPPPPP 978

Qy       1012 PQQPGSSPRGKSRSPAPPADKEAF----------AAEAQKLPGDPPCWTSGLPFPVPPRE 1061
Db        979 PPQPQPPPQQQHQPPPRPVHLQPMQFSTHIQQPPPPQGQQPPHPPPGQQPPPPQPAKPQQ 1038

Qy       1062 VI---------KASP--------------HAPDPSAFS----YAPP------------- 1080
Db       1039 VIQHHHSPRHHKSDPYSTGHLREAPSPLMIHSPQMSQFQSLTHQSPPQQNVQPKKQVTGR 1098

Qy       1081 GHPLPLGLHDTARPVLP----------RPPTISNPPPLI---SSAKHPSVLERQIGAISQ 1127
Db       1099 AGPSPVGQGRGCLPTSPAAVPVPSQELRAASVVQPQPLVVVKEEKIHSPIIR------SE 1152

Qy       1128 GMSVQLH---VPYSEHAKAPV-------GPVTMGLPLPMDPKKLAPFSGV----KQEQLS 1173
Db       1153 PFSPSLRPEPPKHPESIKAPVYVPGPEMKPVDVGRPVIRPPEQNAPPPGAPDKDKQKQ-E 1211

Qy       1174 PRGQAGPPESLGVPTAQEASVLRGTALGSVPGGSITKGIPST-----RVPSDSAITYRGS 1228
Db       1212 PKTPVAPKKDLKIKN----------MGS--WASLVQKHPTTPSSTAKSSSDSFEQFRRA 1258

Qy       1229 ITHGTPADVLYKGTITRIIGEDSPSRLD--RGREDSLPKGHVIYEGKKGHVLSYEGGMSV 1286
Db       1259 AREKEERERKALKAQAEHAEKEKERLRQERMRSRED------------------------ 1293

Qy       1287 TQCSKEDGRSSSGPPHETAAPKRTYDMMEGRVGRAISSASIEGLMGRAIPPERHSPHHLK 1346
Db       1294 -----EDALEQARRAHEEARRRQEQQQQQRQEQQQQQQQQQAAAVAAAATPQAQSSQ---- 1344

Qy       1347 EQHHIRGSITQGIPRSYVEAQEDYLR-REAKLLKRE 1381
Db       1345 ------------PQSMLDQQRELARKREQERRRRE 1367
```

RESULT 82
US-10-210-130-34
; Sequence 34, Application US/10210130
; Publication No. US20040014053A1
; GENERAL INFORMATION:
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Patturajan, Meera
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Miller, Charles E.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Li, Li
; APPLICANT: Berghs, Constance
; APPLICANT: Zhong, Mei
; APPLICANT: Casman, Stacie J.
; APPLICANT: Voss, Edward Z.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Smithson, Glennda
; APPLICANT: Ji, Weizhen
; APPLICANT: Gorman, Linda
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Leite, Mario W.
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Anderson, David W.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Khramtsov, Nikolai V.
; APPLICANT: Ort, Tatiana
; APPLICANT: Ellerman, Karen
; APPLICANT: Rastelli, Luca
; APPLICANT: Agee, Michele L.
; APPLICANT: Chaudhuri, Amitabha
; APPLICANT: Chant, John S.
; APPLICANT: DiPippo, Vincent A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Eisen, Andrew J.
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Giot, Loic
; APPLICANT: Ooi, Chean Eng
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Spaderna, Steven K.
; APPLICANT: Hjalt, Tord
; APPLICANT: Liu, Xiaohong
; APPLICANT: Taupier, Raymond J., Jr.
; APPLICANT: Catterton, Elina
; APPLICANT: Shenoy, Suresh G.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-416C (Cura-716 SMT)
; CURRENT APPLICATION NUMBER: US/10/210,130
; CURRENT FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: 60/309,501
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 60/316,508
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: 60/354,655
; PRIOR FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: 60/310,291
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 60/383,887
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: 60/310,951
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/323,936
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: 60/381,039
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/311,292
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/311,979
; PRIOR FILING DATE: 2001-08-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 34
; LENGTH: 3186
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-210-130-34

Query Match 2.8%; Score 374.5; DB 15; Length 3186;
Best Local Similarity 19.3%; Pred. No. 1.9e-08;
Matches 430; Conservative 282; Mismatches 762; Indels 759; Gaps 108;

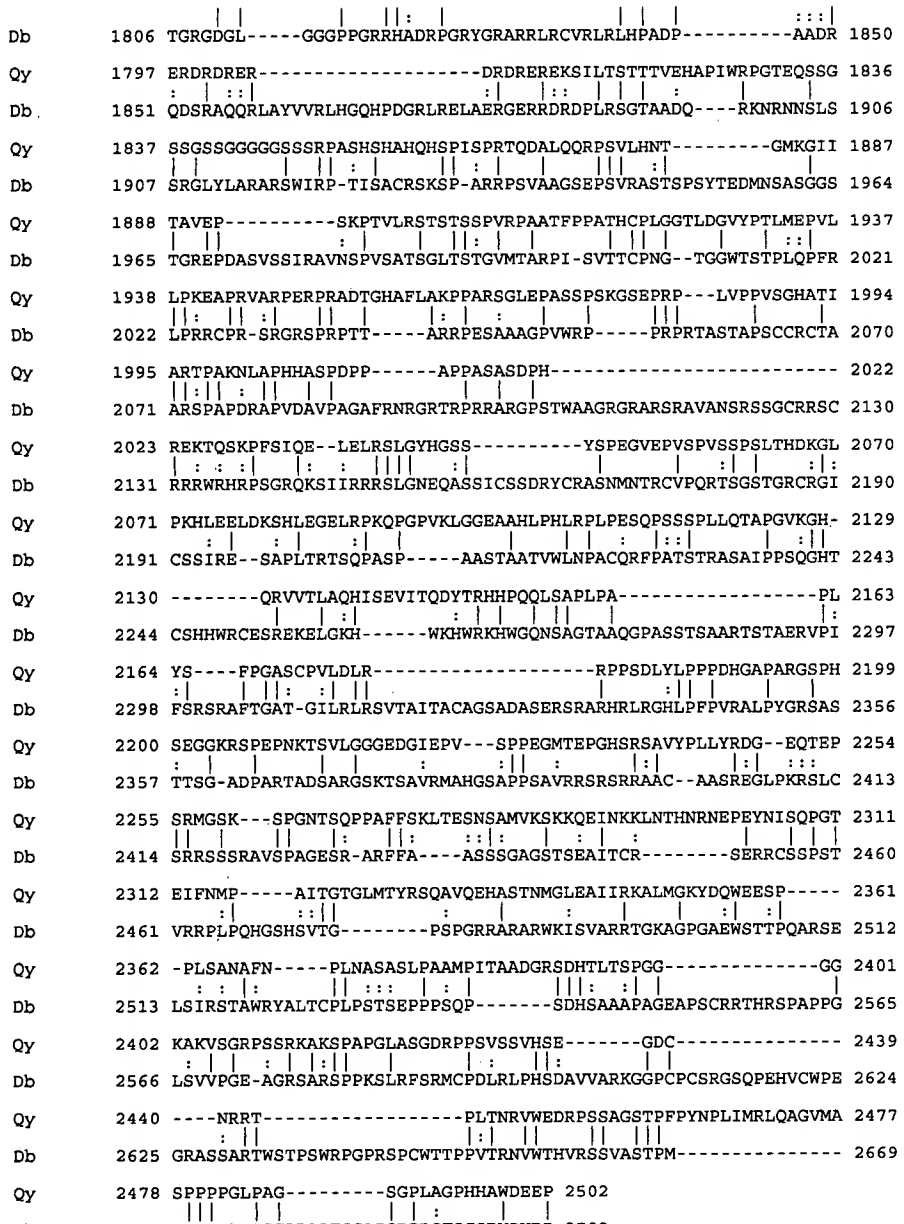
```
Qy         78 RSQELHLRPESHSYLPE---LGKSEMEFIESKRPRLELLPDPLLRPSPLLATGQPAGSED 134
Db       1292 RSLDMAIDPDSLTYTVQHQVLGSDQWTALVTG----------LREPGWAATGLRKGVQH 1340

Qy        135 LTKDRSLTGKLEPVSPPSPP-------HTDPELELVPPRLSKEELIQNMDRVDREITMVE 187
Db       1341 IFRVLSTTVKSS--SKPSPPSEPVQLLEHGPTLEEAPAMLDKPDIVY----------VVE 1388

Qy        188 QQISKLKKKQQQLEEEAAKPPEPEKPVSPPPIESKHRSLVQIIYDENRKKAEAAHRILEG 247
```





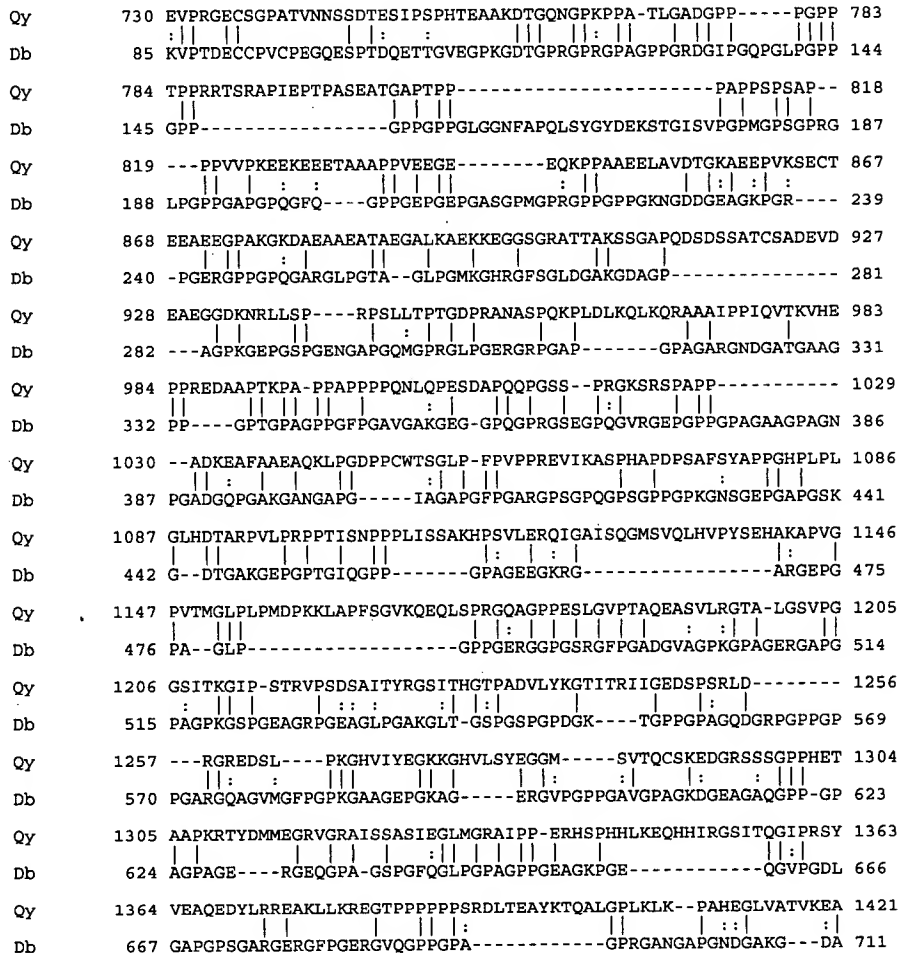
RESULT 83
US-10-084-846A-3
; Sequence 3, Application US/10084846A

```
; Publication No. US20040006026A1
; GENERAL INFORMATION:
;  APPLICANT: WEITNAUER, GABRIELE
;  APPLICANT:  MUHLENWEG, AGNES
;  APPLICANT:  TREFZER, AXEL
;  APPLICANT:  BECHTHOLD, ANDREAS
;  TITLE OF INVENTION: AVILAMYCIN DERIVATIVES
;  FILE REFERENCE: 1974-005
;  CURRENT APPLICATION NUMBER: US/10/084,846A
;  CURRENT FILING DATE:  2003-02-25
;  PRIOR APPLICATION NUMBER: PCT/EP01/09815
;  PRIOR FILING DATE: 2001-08-24
;  PRIOR APPLICATION NUMBER: DE 101 09 166.4
;  PRIOR FILING DATE: 2001-02-25
;  NUMBER OF SEQ ID NOS: 120
;  SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 3
;   LENGTH: 19695
;   TYPE: PRT
;   ORGANISM: Streptomyces viridochromogenes
;   FEATURE:
;   OTHER INFORMATION: Protein 1: amino acid sequence encoded by coding strand 1.
;   OTHER INFORMATION: Start codon: gga, Start position: nucleotide 1.
US-10-084-846A-3

  Query Match          2.8%;  Score 374;  DB 15;  Length 19695;
  Best Local Similarity   19.4%;  Pred. No. 1.7e-07;
  Matches  531;  Conservative  247;  Mismatches  948;  Indels 1008;  Gaps  126;

Qy        485 RSYRRRGKS------QQQQQQQQQQQQQQQQQPMPRSSQEEKDEKEKEKEAEKEEEKPEV 538
              | : |||:       ::|    |:|    |::   |    : |  |:|     ||:   :|
Db        261 RYHHRRGRQVAPAHPHRRQDAHGQEQHAVQRERRPEQQRVEVAEQEHHGAAEQAAAQP-- 318

Qy        539 ENDKEDLLKEKTDDTSGEDNDEKEA--VASKGRKTANSQGRRKGRITRSMA--------- 587
                              |:    | |  :   ||  |  :  |:||  |: |
Db        319 ----------------GQPGVGKPAQGLGRGGRPGAPYRRGRRGRTGRTRAALGDGPGGA 362

Qy        588 ------------------------------------NE---ANSEEAITPQQSAELA 605
                                                   ||    |:     || ::|  |
Db        363 GGPGRLGGPGGPGGPAVPARTAPAARETAVAARRTAPAWNESAPASLHTVFTPPRTAAPA 422

Qy        606 SMELNESSRWTEEEMETAKKGLLEHGR------NWSAIARMVGSKTVSQCKNFYFNYK-- 657
              : :   : |    :          ||      :|: : |::          |:  :   :
Db        423 ASDTPPTPRRLRRQRRRRSVSRFVPGRSRPVWLHWTNLGRLM-------CRIAHQTRRGA 475

Qy        658 ---------------KRQNLDEILQQHKLKMEKERNARRKKKKAPAAASEEAAFPPVVED 702
                            :|: :|  :: | |      : | || :|||          | ::
Db        476 DLGHVFDTAPGPGPARRRLVDRPVRLHHL------HRRSKKDRAPA--------QPALKS 521

Qy        703 EEMEASGVSGNEEEMVEEAEALHASGNEVPRGECSGPATVNNSSDTESIPSPHTEAAKDT 762
              :              | |      | : | |    | | :             | |    : |
Db        522 SHVR------------EVATCRSRSESSVCRMSASRPCST----------PPRTTCWRPT 559

Qy        763 GQNGPKPPATLGADGPPPGPPTPPRRTSRAPIEPTPASEATGAPTPPPAPPSPSAPPPVV 822
              ::     | ::     |    |: ||          :| | ::  |:     |:  | ||
Db        560 TRSPRSSPTSVSWASPTHASPSWPR--------SSPPSASSRRPSTSSTSPASCAAPPRG 611

Qy        823 PKEEKEEETAAAPPVEEGEEQKPPAAEELAVDTGKAEEPVKS-ECTEEAEEGPAKGKDAE 881
                        :  |          |      :: | ::     |    |  :     : : :
Db        612 RAWATSSSRTSVSPTRSARSSAPSRTRTSSMSTARSRPRTTSRRSTPSSSRTSRRSRRSC 671

Qy        882 AAEATAEGALKAEKKEGGSGR----------ATTAKSSGAPQDSDSSATCSADEVDEAEG 931
                |      |: :  ::    | |           ::   || ||  : |  ||:      :
Db        672 PASRRTRGSRRTSRRRRPSRRPRRSWRGATPSSPTASSRAPSATSSCTTCTCSPPSPSST 731

Qy        932 GDKNRLLS---------------PRPSLLTP---------------TGDPRANASPQKP 960
                  :   |                ||||  ||                :  | |  ||  |
Db        732 SSTSTRTSSTRTSRTSSAPWSPPPRPSSSTPSWRPTSPSSTTTRPSNSSSPSARTSPASP 791

Qy        961 LDLKQLKQRAAAIPPIQVTKVHEPPREDAAPTKPAPPAPPPPQNLQPESDAPQQPGSSPR 1020
                        |:|  |    :   | | : ||  |:  |||||:   | |  |    :| |
Db        792 ------PSPASASTP-SACRPTRPARRNPAP-GPSRRAPPPPR--PPASSTPTSRRASSR 841

Qy       1021 GKSRSPAP----PADKEAFAAEAQKL---------------------------------- 1042
               :| ||:|    ||     |   :::
Db        842 RRS-SPSPTWWRPAPSPRPARRGRRVWRARTMCRTGTWSSSASTCRRNSTTTLLSNAQVR 900

Qy       1043 ---------PGDPPCWTS----------GLPFPVPPREVI----KASPHAPDPSAFSYAP 1079
                      |  |||               ||  |    ::    | | |   | |    |
Db        901 RGRLLQGAAPCCFPCWDGAGKADHPSVFALPEDVKSGTLLGTTGKGSNHVDKPRA---KP 957

Qy       1080 PGH------------------------PLPLGLHDTARPVLPRPPTISNPPPLISSAKH 1114
               |                          |  ||   || ::  |  :|   |  :
Db        958 QGSLSADQEGQRQSRGHRDRLQARQRRTRLGRGLCSAARGLV--PAALSGERPSTAQGVR 1015

Qy       1115 PSVLERQ-IGAISQ---------GMSVQLHV------PYSEHAKAPVGPVTMGLPLPMDP 1158
                  || :|| :          |  |:|         |:  |              |
Db       1016 ERPCPRQCVGAGADRSGFGGRWFGCRVRLVFEDQGWDDYTSWLK-------------NDR 1062

Qy       1159 KKLA------------PFSGVKQEQLSPRGQAGPPESLG-------------------VP 1187
               | ||            ||:|:           | || |                    :
Db       1063 KMLARINKLIEDVRRDPFTGI----------GKPEPLKYHLPGAWSRRIDDEHRLVYLV 1111

Qy       1188 TAQEASVL-------------RGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHGTP 1234
               | :|  :|             |  ||  |    |:      : | || |:|    |
Db       1112 TDKEIVILAARYHYLPRWSSRAGMVLGWAPR---TRCFRRWELCSCSASTFRRRST---- 1164

Qy       1235 ADVLYKGTITRIIGEDSPSRLDRGREDS------LPKGHVIYEGKKGHVLSYEGGMSVTQ 1288
                                  | |   | : |       |     :|    |:   :| :
Db       1165 --------------TSSCMRRSEGSDSSESGPWLFPVPPADWELLSHHLPIGQGPVQSAS 1210

Qy       1289 CSKEDG----------RSSSGPP-----HETAAPKRTYDMMEGR-------VGR---AIS 1323
                : :|           :|:  |       : | |::   :  ||        | |    ||
Db       1211 ADQLEGIIWRSVPVRISASAWVPKRARVKQWAPPEKQNSLAGGRSPMRSRNVARRPTPIS 1270

Qy       1324 SASIEGLMGRAIPPERHSPHHLKEQHH----IRGSITQGIPRS------------YVEAQ 1367
                |         |:  |||||   :  |     :  |  : : ||            :  |
Db       1271 QA--------AVRRERHSPR--RASHSAWVIVAASSMRRVKRSAWGHPGRRTRPLFRSRQ 1320

Qy       1368 EDYLRREAKLLKREGTPP----------------------------PP-----PPSRDL 1393
                | |: |: |  |||                              ||       |||
Db       1321 RG--RSESLLVTRSITPPLSSECWNGIGHLSGYRSVRSIQVGATKMLPPTGASEASSRDQ 1378

Qy       1394 TEAYKTQ-ALGPL----KLKPAHE-----------------------GLVATVKEAGRS 1424
                |: : | |    : :|  |                        |: :|| ||
Db       1379 VWDSSTRCSQGALQSSSQSRPVPEMPYVRSASTRQNREWRITWAWSARSLMISVKWAGAE 1438

Qy       1425 IHEIPREELRHTPELPLAPRPLKEGSITQGT-----PLKYD---------TGASTTGSKK 1470
                : |           :|    |    :|:     ||:             ||:    :
Db       1439 FGQGP-------APIPRWGRQRPAGRSGRGSSVGLLPLRRPRWRSAGRSCPGAARRSGLR 1491

Qy       1471 HDVRSLIGSPGRTFPPVHPLDVMADARALERACYEESL----------KSRPG-TASSSG 1519
               || |    : ||              |    ||    |              : | | |: | |
Db       1492 HD-RCRPAARGRRM----------RRRRRRAAGRGSRCGGSWCAGWPRQRWGATSRAPG 1540

Qy       1520 GSIARGAPVIVPELGKPRQS----------PLTYEDHGAPFAGHLPRGSPVTMREPTP 1567
               | | |    :| | : | |            |       |  ||      |  | |:|
Db       1541 GRRATGRGPRLPGLRRGRVSGRRGARRGRAARPRPAGGRTADRAGRRAH-SRATSRQPPG 1599

Qy       1568 RLQEGSLS------SSKASQDRKLT----------STPREI------AKSPHSTVPEHH 1604
               | | |: |      | :|  : |:|          | |        |: |
Db       1600 RPQGGARSHFVSARSGRAPSNGKITWRADAGGSADHLTQRSFPAVAGTARQPQQAHARRV 1659

Qy       1605 PHPISPYEHLLRGVSGVDLYRSHIPLAFDPTSIPRGI---------------PLDAAA- 1647
               |   || : | ||  | | |:| ||      ::  |               | | |
Db       1660 PGDGSPPD-LARG--GARLARTHAPLRRTAGAVEVGAGASSYGADGADRDGPGPADGAGG 1716

Qy       1648 -------------AYYLPRHLAPNPTYPHLYPPYLIRGYPDTAALENRQTIINDYITSQ 1693
                          |   ||   |:|  |   |   :|  |   : |:||
Db       1717 EERGPARGRRRRGDARRRPR---PDPG-PGGLPGRTVR--PADQSGEHRQ---------- 1760

Qy       1694 QMHHNTATAMAQ-RADMLRGLSPRESSLALNYAAGPRGIIDLSQVPH-LPVLVPPTPGTP 1751
                ||:      : |||  || : |     |    ||        | | | ||    ||
Db       1761 --HHSLGGCPGRARAD--RGGARRR---RLRPEIGP--------VEHRLAVLPARRPGAV 1805

Qy       1752 ATAMDRLAYLPTAPQPFSSRHSSSP--------------LSPGGPTHLTKPTTTSSSER 1796
```

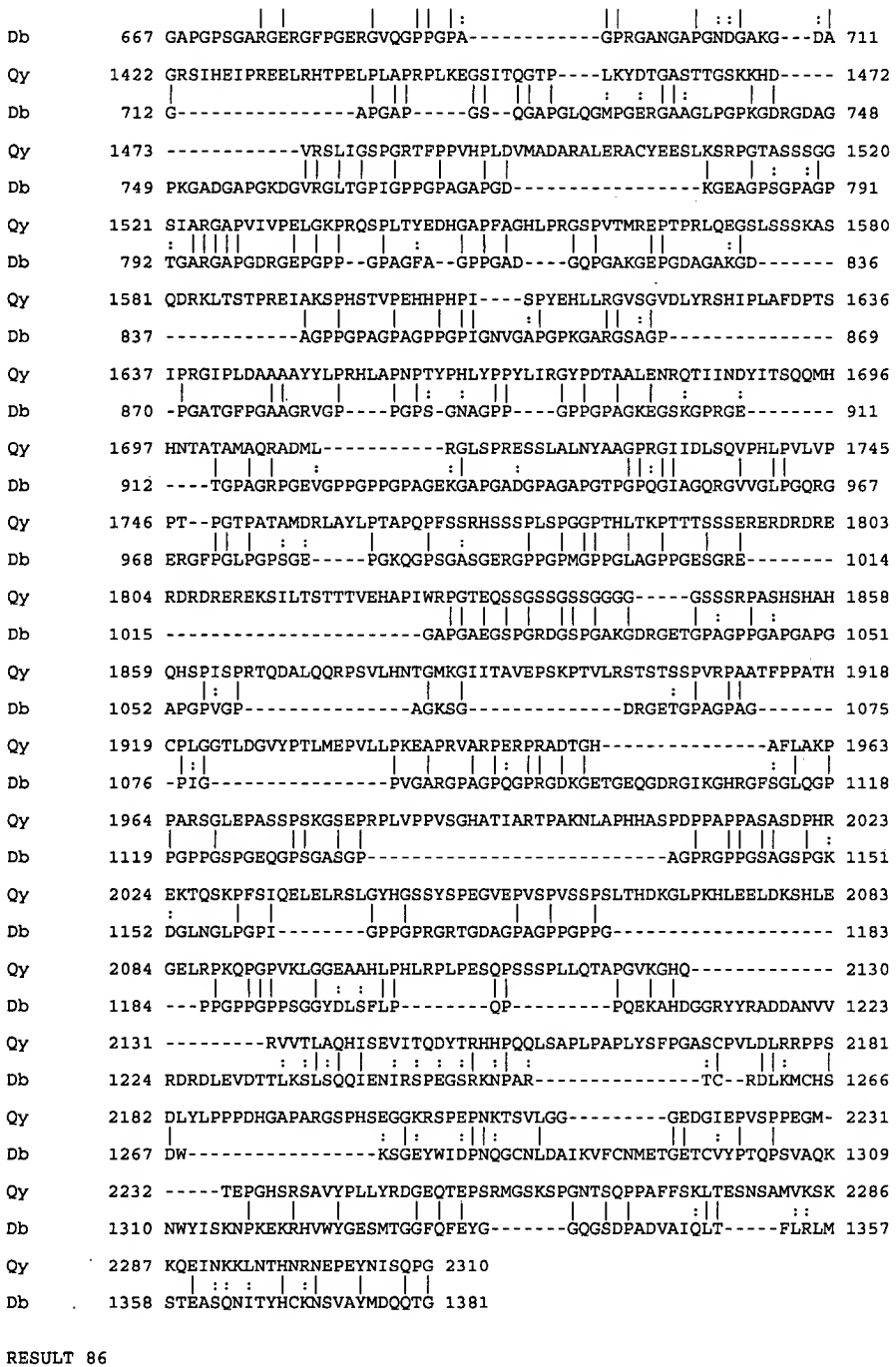


RESULT 84
US-10-402-089-2
; Sequence 2, Application US/10402089
; Publication No. US20040005663A1
; GENERAL INFORMATION:
; APPLICANT: Bell, Marcum P.
; APPLICANT: Neff, Thomas B.
; APPLICANT: Polarek, James W.
; APPLICANT: Seeley, Todd W.
; TITLE OF INVENTION: PORCINE COLLAGENS AND GELATINS
; FILE REFERENCE: FP0402.3 CON
; CURRENT APPLICATION NUMBER: US/10/402,089
; CURRENT FILING DATE: 2003-03-26
; PRIOR APPLICATION NUMBER: US 09/709,700
; PRIOR FILING DATE: 2000-11-10
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 1463
; TYPE: PRT
; ORGANISM: Bos Taurus
US-10-402-089-2

Query Match 2.8%; Score 371; DB 15; Length 1463;
Best Local Similarity 21.0%; Pred. No. 1.1e-08;
Matches 371; Conservative 97; Mismatches 646; Indels 650; Gaps 79;



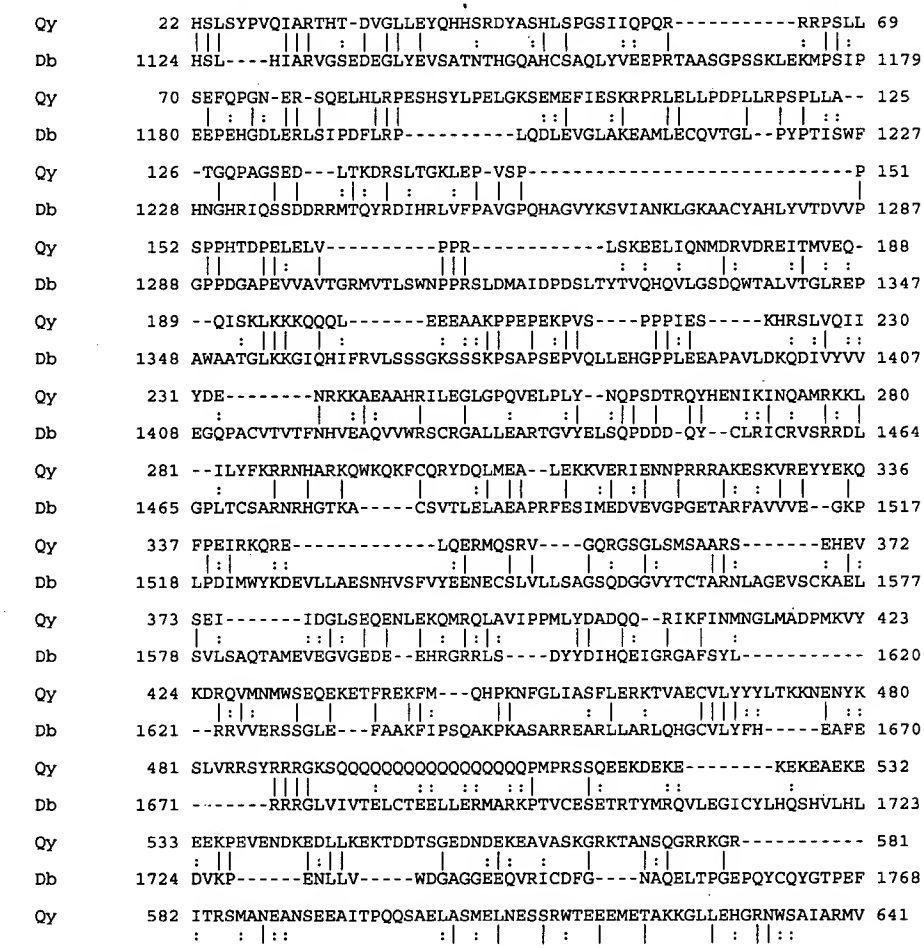
```
Qy       1422 GRSIHEIPREELRHTPELPLAPRPLKEGSITQGTP----LKYDTGASTTGSKKHD----- 1472
Db        712 G----------------APGAP-----GS--QGAPGLQGMPGERGAAGLPGPKGDRGDAG 748

Qy       1473 ------------VRSLIGSPGRTFPPVHPLDVMADARALERACYEESLKSRPGTASSSGG 1520
Db        749 PKGADGAPGKDGVRGLTGPIGPPGPAGAPGD----------------KGEAGPSGPAGP 791

Qy       1521 SIARGAPVIVPELGKPRQSPLTYEDHGAPFAGHLPRGSPVTMREPTPRLQEGSLSSSKAS 1580
Db        792 TGARGAPGDRGEPGPP--GPAGFA--GPPGAD----GQPGAKGEPGDAGAKGD------- 836

Qy       1581 QDRKLTSTPREIAKSPHSTVPEHHPHPI----SPYEHLLRGVSGVDLYRSHIPLAFDPTS 1636
Db        837 ------------AGPPGPAGPAGPPGPIGNVGAPGPKGARGSAGP-------------- 869

Qy       1637 IPRGIPLDAAAAYYLPRHLAPNPTYPHLYPPYLIRGYPDTAALENRQTIINDYITSQQMH 1696
Db        870 -PGATGFPGAAGRVGP----PGPS-GNAGPP----GPPGPAGKEGSKGPRGE-------- 911

Qy       1697 HNTATAMAQRADML-----------RGLSPRESSLALNYAAGPRGIIDLSQVPHLPVLVP 1745
Db        912 ----TGPAGRPGEVGPPGPPGPAGEKGAPGADGPAGAPGTPGPQGIAGQRGVVGLPGQRG 967

Qy       1746 PT--PGTPATAMDRLAYLPTAPQPFSSRHSSSPLSPGGPTHLTKPTTTSSSERERDRDRE 1803
Db        968 ERGFPGLPGPSGE-----PGKQGPSGASGERGPPGPMGPPGLAGPPGESGRE-------- 1014

Qy       1804 RDRDREREKSILTSTTTVEHAPIWRPGTEQSSGSSGSSGGGG-----GSSSRPASHSHAH 1858
Db       1015 ----------------------GAPGAEGSPGRDGSPGAKGDRGETGPAGPPGAPGAPG 1051

Qy       1859 QHSPISPRTQDALQQRPSVLHNTGMKGIITAVEPSKPTVLRSTSTSSPVRPAATFPPATH 1918
Db       1052 APGPVGP---------------AGKSG--------------DRGETGPAGPAG------- 1075

Qy       1919 CPLGGTLDGVYPTLMEPVLLPKEAPRVARPERPRADTGH---------------AFLAKP 1963
Db       1076 -PIG----------------PVGARGPAGPQGPRGDKGETGEQGDRGIKGHRGFSGLQGP 1118

Qy       1964 PARSGLEPASSPSKGSEPRPLVPPVSGHATIARTPAKNLAPHHASPDPPAPPASASDPHR 2023
Db       1119 PGPPGSPGEQGPSGASGP--------------------------AGPRGPPGSAGSPGK 1151

Qy       2024 EKTQSKPFSIQELELRSLGYHGSSYSPEGVEPVSPVSSPSLTHDKGLPKHLEELDKSHLE 2083
Db       1152 DGLNGLPGPI--------GPPGPRGRTGDAGPAGPPGPPG-------------------- 1183

Qy       2084 GELRPKQPGPVKLGGEAAHLPHLRPLPESQPSSSPLLQTAPGVKGHQ------------- 2130
Db       1184 ---PPGPPGPPSGGYDLSFLP--------QP---------PQEKAHDGGRYYRADDANVV 1223

Qy       2131 ---------RVVTLAQHISEVITQDYTRHHPQQLSAPLPAPLYSFPGASCPVLDLRRPPS 2181
Db       1224 RDRDLEVDTTLKSLSQQIENIRSPEGSRKNPAR---------------TC--RDLKMCHS 1266

Qy       2182 DLYLPPPDHGAPARGSPHSEGGKRSPEPNKTSVLGG---------GEDGIEPVSPPEGM- 2231
Db       1267 DW-----------------KSGEYWIDPNQGCNLDAIKVFCNMETGETCVYPTQPSVAQK 1309

Qy       2232 -----TEPGHSRSAVYPLLYRDGEQTEPSRMGSKSPGNTSQPPAFFSKLTESNSAMVKSK 2286
Db       1310 NWYISKNPKEKRHVWYGESMTGGFQFEYG-------GQGSDPADVAIQLT-----FLRLM 1357

Qy       2287 KQEINKKLNTHNRNEPEYNISQPG 2310
Db       1358 STEASQNITYHCKNSVAYMDQQTG 1381

RESULT 85
US-10-402-072A-2
; Sequence 2, Application US/10402072A
; Publication No. US20040018592A1
; GENERAL INFORMATION:
;  APPLICANT: Bell, Marcum  P.
;  APPLICANT:  Neff, Thomas B.
;  APPLICANT:  Polarek, James W.
;  APPLICANT:  Seeley, Todd W.
;  TITLE OF INVENTION: BOVINE COLLAGENS AND GELATINS
;  FILE REFERENCE: FP0402.2 CON
;  CURRENT APPLICATION NUMBER: US/10/402,072A
;  CURRENT FILING DATE:  2003-03-26
;  PRIOR APPLICATION NUMBER: US 09/709,700
;  PRIOR FILING DATE: 2000-11-10
;  NUMBER OF SEQ ID NOS: 72
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
;   LENGTH: 1463
;   TYPE: PRT
;   ORGANISM: Bos Taurus
US-10-402-072A-2

  Query Match             2.8%;  Score 371;  DB 15;  Length 1463;
  Best Local Similarity   21.0%;  Pred. No. 1.1e-08;
  Matches  371;  Conservative   97;  Mismatches  646;  Indels  650;  Gaps   79;

Qy        730 EVPRGECSGPATVNNSSDTESIPSPHTEAAKDTGQNGPKPPA-TLGADGPP-----PGPP 783
Db         85 KVPTDECCPVCPEGQESPTDQETTGVEGPKGDTGPRGPRGPAGPPGRDGIPGQPGLPGPP 144

Qy        784 TPPRRTSRAPIEPTPASEATGAPTPP----------------------PAPPSPSAP-- 818
Db        145 GPP----------------GPPGPPGLGGNFAPQLSYGYDEKSTGISVPGPMGPSGPRG 187

Qy        819 ---PPVVPKEEKEEETAAAPPVEEGE--------EQKPPAAEELAVDTGKAEEPVKSECT 867
Db        188 LPGPPGAPGPQGFQ----GPPGEPGEPGASGPMGPRGPPGPPGKNGDDGEAGKPGR---- 239

Qy        868 EEAEEGPAKGKDAEAAEATAEGALKAEKKEGGSGRATTAKSSGAPQDSDSSATCSADEVD 927
Db        240 -PGERGPPGPQGARGLPGTA--GLPGMKGHRGFSGLDGAKGDAGP---------------- 281

Qy        928 EAEGGDKNRLLSP----RPSLLTPTGDPRANASPQKPLDLKQLKQRAAAIPPIQVTKVHE 983
Db        282 ---AGPKGEPGSPGENGAPGQMGPRGLPGERGRPGAP-------GPAGARGNDGATGAAG 331

Qy        984 PPREDAAPTKPA-PPAPPPPQNLQPESDAPQQPGSS--PRGKSRSPAPP----------- 1029
Db        332 PP----GPTGPAGPPGPPGAVGAKGEG-GPQGPRGSEGPQGVRGEPGPPGPAGAAGPAGN 386

Qy       1030 --ADKEAFAAEAQKLPGDPPCWTSGLP-FPVPPREVIKASPHAPDPSAFSYAPPGHPLPL 1086
Db        387 PGADGQPGAKGANGAPG-----IAGAPGFPGARGPSGPQGPSGPPGPKGNSGEPGAPGSK 441

Qy       1087 GLHDTARPVLPRPPTISNPPPLISSAKHPSVLERQIGAISQGMSVQLHVPYSEHAKAPVG 1146
Db        442 G--DTGAKGEPGPTGIQGPP-------GPAGEEGKRG-----------------ARGEPG 475

Qy       1147 PVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPESLGVPTAQEASVLRGTA-LGSVPG 1205
Db        476 PA--GLP------------------GPPGERGGPGSRGFPGADGVAGPKGPAGERGAPG 514

Qy       1206 GSITKGIP-STRVPSDSAITYRGSITHGTPADVLYKGTITRIIGEDSPSRLD-------- 1256
Db        515 PAGPKGSPGEAGRPGEAGLPGAKGLT-GSPGSPGPDGK----TGPPGPAGQDGRPGPPGP 569

Qy       1257 ---RGREDSL----PKGHVIYEGKKGHVLSYEGGM-----SVTQCSKEDGRSSSGPPHET 1304
Db        570 PGARGQAGVMGFPGPKGAAGEPGKAG-----ERGVPGPPGAVGPAGKDGEAGAQGPP-GP 623

Qy       1305 AAPKRTYDMMEGRVGRAISSASIEGLMGRAIPP-ERHSPHHLKEQHHIRGSITQGIPRSY 1363
Db        624 AGPAGE----RGEQGPA-GSPGFQGLPGPAGPPGEAGKPGE-----------QGVPGDL 666

Qy       1364 VEAQEDYLRREAKLLKREGTPPPPPSRDLTEAYKTQALGPLKLK--PAHEGLVATVKEA 1421
```



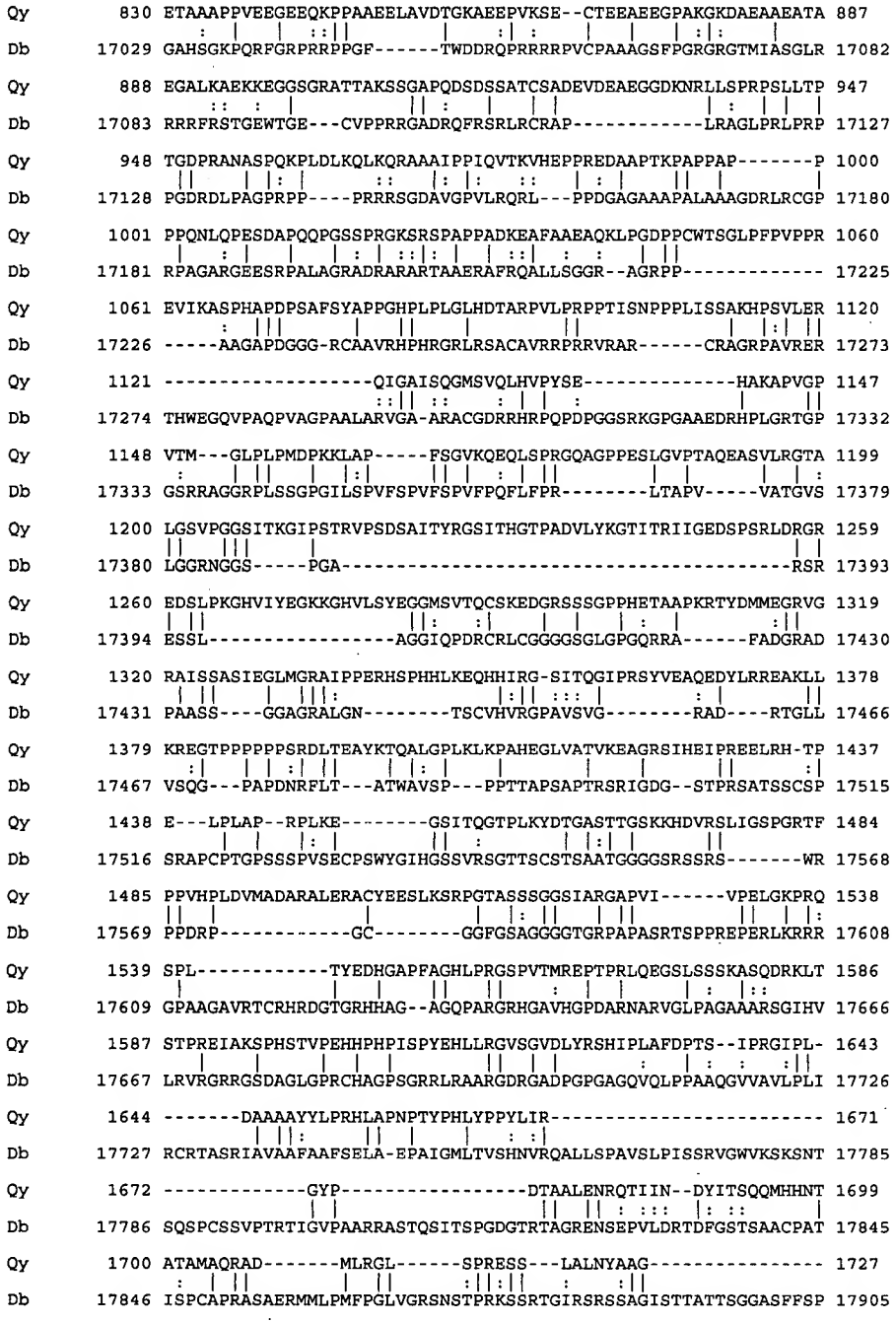
RESULT 86

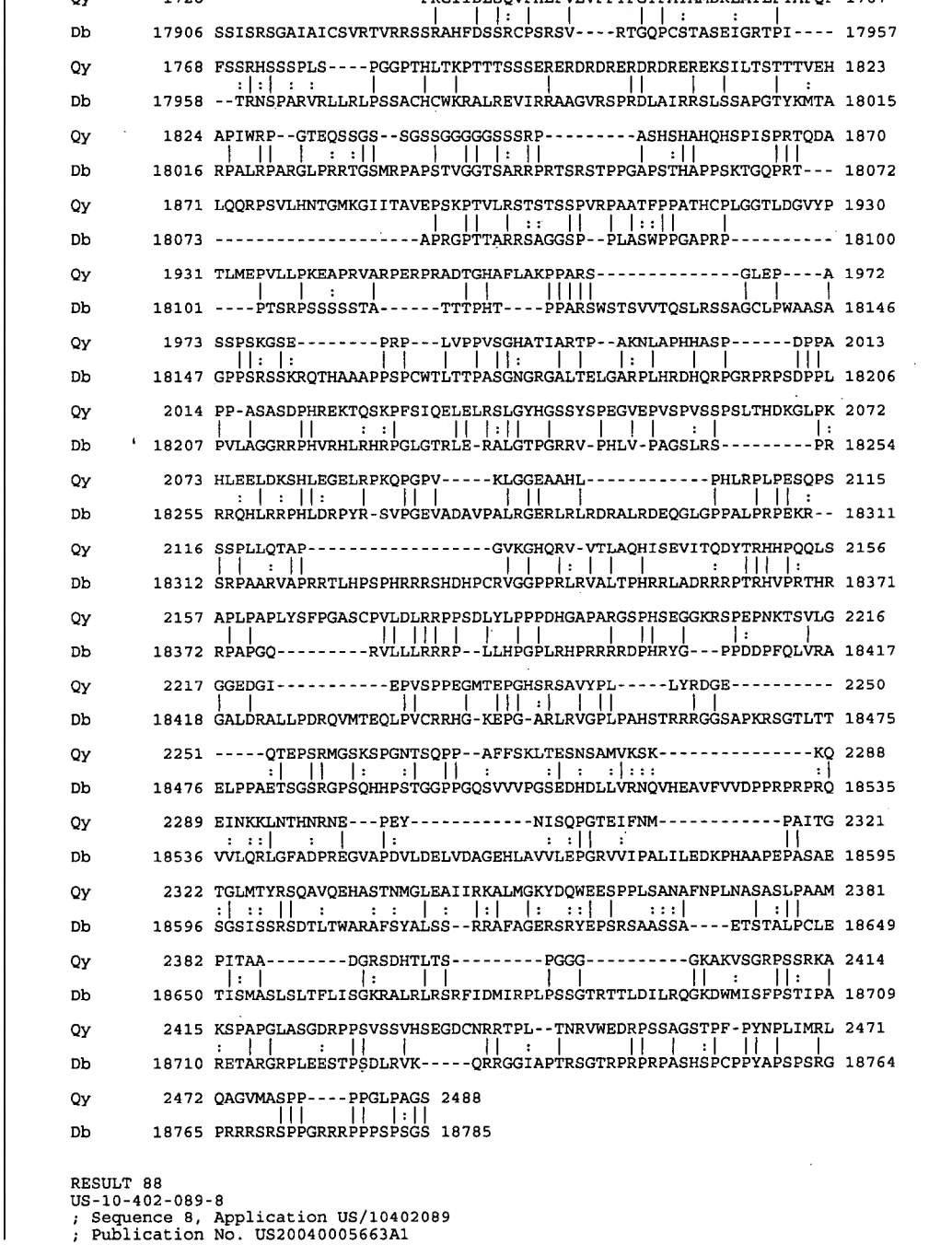
US-10-379-381-4
; Sequence 4, Application US/10379381
; Publication No. US20030228595A1
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001360
; CURRENT APPLICATION NUMBER: US/10/379,381
; CURRENT FILING DATE: 2003-03-05
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 3262
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-379-381-4

Query Match 2.8%; Score 369.5; DB 15; Length 3262;
Best Local Similarity 20.6%; Pred. No. 3.3e-08;
Matches 434; Conservative 273; Mismatches 784; Indels 619; Gaps 113;



```
Db      1769 VAPEIVNQSPVSGVTDIWPVGVVAFLCLTGISPFVGENDRTT----LMNIRNYNVAFEET 1824

Qy       642 GSKTVS-QCKNFYFNY----KKRQNLDEILQQHKLKME----------KERNARRKKKK 685

Db      1825 TFLSLSREARGFLIKVLVQDRLRPTAEETLEHPWFKTEAKGAEVSTDHLKLFLSRRRWQR 1884

Qy       686 APAAASEEAAFPPVVE-------------DEEMEASGVSGNEEEMVEEAEALHA----- 726

Db      1885 SQISYKCHLVLRPIPELLRAPPERVWVAMPRRQPPSGGLSSSSDSEEEELEELPSVPRPL 1944

Qy       727 ----SGNEVPRGEC----------SGPATVNNSSDTESIPS-----PHTEAAKDTGQNG 766

Db      1945 QPEFSGSRVSLTDIPTEDEALGTPEAGAATPMDWQEQERTPSKDQEAPSPEALPSPGQES 2004

Qy       767 PKPPATLGADGPPPGPPTPPRRTSRAPIEPTPAS--------EATGAPTPPPAPPSPSA- 817

Db      2005 P--------DGPSPRRPELRRGSSAESALPRVGSREPGRSLHKAASVELPQRRSPSPGAT 2056

Qy       818 --------PPPVVPKEEKEEETAAAPPVEEGEEQ--KPPAAEELAVDTGKAEEP--VKSE 865

Db      2057 RLTRGGLGEGEYAQRLQALRQRLLRGGPEDGKVSGLRGPLLESLG---GRARDPRMARAA 2113

Qy       866 CTEEA--EEGPAKGKDAEAAEATAEGALKAEKKEGGSGRATTAKSSGAPQDSDSSATCSA 923

Db      2114 SSEAAPHHQPPPESRGLQKSSSFSQG--EAEPR----GR---HRRAGAPL---------- 2154

Qy       924 DEVDEAEGGDKNRLLSPRPSLLTPTGDPRANASPQKPLDLKQLKQRAAAIPPIQVTKVHE 983

Db      2155 -EIPVARLGARRLQESPSLSALSETQPP----SPARP----------SVPKLSITKSPE 2198

Qy       984 PPREDAAPTKPAPPAPPPPQNL-------QPESDAPQQPGSSPRGKSRSPAPPADKEAFA 1036

Db      2199 P----SAVTSRDSPQPPEPQPVPEKVPEPKPEPVRAAKPAQPPLA-LQMPTQPLTPYAQI 2253

Qy      1037 AEAQKL------PGDPPCWTSGLPFPVPPREVIKASPHAPDPSAFSYAPPGHPLPLGLHD 1090

Db      2254 MQSLQLSSPTLSPQDP---------AVPPSE--------PKPHAAVFARVASP-PPGVSE 2295

Qy      1091 TARPVLPRPPTISNPPPLISSAKHPSVLERQIGAISQGMSVQLH---VPYSEHAKAPVGP 1147

Db      2296 ------KRVPSARTPPVLAEKARVPTVPPRP-GSSLSGSIENLESEAVFEAKFKRSRESP 2348

Qy      1148 VTMGLPL--PMDPKKLAPFSGVKQEQLSPRGQAGPPESLGVPTAQEASVLRGTALGSVPG 1205

Db      2349 LSRGLRLLSRSRSEERGPFRGAEDDGIYRPSPAGTPLEL-VRRPERSRSVQDLRVAGEPG 2407

Qy      1206 ----------GSITKGIPSTRVPS------DSAITYRGSITHGTPADVLYK---GTITRI 1246

Db      2408 LVRRLSLSLSQKLRRTPPGQRHPAWESRSGDGESSEGGSSARASPVLAVRRRLSSTLERL 2467

Qy      1247 IGEDSPSRLDR--GREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPHET 1304

Db      2468 -----SSRLQRSGSSEDS--------GGASGRSTPLFGRLR-RATSEGESLRRLGVPH-- 2511

Qy      1305 AAPKRTYDMMEGRVGRAISSA-----SIEGLMGRAIPPERHSPH-----HLKE-----QH 1349

Db      2512 -------NQLGSQTGATTPSAESLGSEASGTSGSSAPGESRSRHRWGLSRLRKDKGLSQP 2564

Qy      1350 HIRGSITQGIPRSYVEAQEDY------LRREAKLLKREGTP----PPPPSRDLTEAYKT 1399

Db      2565 NLSSSVQEDLGHQYVPSESDFPPVFHIKLKDQVLLEGEAATLLCLPAACPAPRISWMKDK 2624

Qy      1400 QALGPLKLKPAHEGLVATVKEAGRSIHEIPREELRHTPELPLAPRPLKE-------GSIT 1452

Db      2625 QS---LRSEPS---VVIVSCKDGRQLLSIPRAGKRHA--------GLYECSATNVLGSIT 2670

Qy      1453 QG---------------TPLKYDTGASTTGSKKHDVRSLIGSPG-RTFPPVHPLDVMAD 1495

Db      2671 SSCTVAVARIPGKLAPPEVPQTY----------HDTALVVWKPGDGRAPCTYTLERRVD 2719

Qy      1496 ARALERA-------CY--------------EESLKSRPGTASSSGGS---IARGAPVIVP 1531

Db      2720 GESVWHPVSSGIPDCYYNVTQLPVGVTVRFRVACSNRAGQGPFSNPSEKVFIRGTPDSPA 2779

Qy      1532 E-LGKPRQSPLTYEDHGAPFAGHLPRGSPVTMREPTPRL-----QEGSLSSSKASQDRKL 1585

Db      2780 QPAAAPRDAPVTSGPTRAP-----PPDSPTSL-APTPALAPPASQASTLSPSTSSMSANQ 2833

Qy      1586 TSTPREIAKSPHSTVPEHH----------PHP----ISPYEHLLRGVSGVDLYRSHIP-- 1629

Db      2834 ALSSLKAVGPPPATPPRKHRGLLATQQAEPSPPSIVVTPSEP----------RSFVPDT 2882

Qy      1630 LAFDPTSIPRGI-PLDAAAAYYL----------PRHLAPNP-------TYPHLYPPYLIR 1671

Db      2883 GTLTPTSSPQGVKPAPSSTSLYMVTSFVSAPPAPQAPAPEPPPEPTKVTVRSLSPAKEVV 2942

Qy      1672 GYPDTAALENRQTIINDYITSQQMHHNTATAMAQRAD-MLRG--LSPRESSLALNYAA-G 1727

Db      2943 SSPTPESTTLRQGPLRNPTPSWRRRPGGALALCGHAGRMLRAERLSPR----FVPYAAEG 2998

Qy      1728 PRGIIDLSQV 1737

Db      2999 KRRVLQEYEV 3008


RESULT 87
US-10-084-846A-8
; Sequence 8, Application US/10084846A
; Publication No. US20040006026A1
; GENERAL INFORMATION:
;  APPLICANT: WEITNAUER, GABRIELE
;  APPLICANT:  MUHLENWEG, AGNES
;  APPLICANT:  TREFZER, AXEL
;  APPLICANT:  BECHTHOLD, ANDREAS
;  TITLE OF INVENTION: AVILAMYCIN DERIVATIVES
;  FILE REFERENCE: 1974-005
;  CURRENT APPLICATION NUMBER: US/10/084,846A
;  CURRENT FILING DATE:  2003-02-25
;  PRIOR APPLICATION NUMBER: PCT/EP01/09815
;  PRIOR FILING DATE: 2001-08-24
;  PRIOR APPLICATION NUMBER: DE 101 09 166.4
;  PRIOR FILING DATE: 2001-02-25
;  NUMBER OF SEQ ID NOS: 120
;  SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 8
;   LENGTH: 19608
;   TYPE: PRT
;   ORGANISM: Streptomyces viridochromogenes
;   FEATURE:
;   OTHER INFORMATION: Protein 3: amino acid sequence encoded by coding strand 1.
;   OTHER INFORMATION: Start codon: atc, Start position: nucleotide 3.
US-10-084-846A-8

  Query Match             2.8%;  Score 367;  DB 15;  Length 19608;
  Best Local Similarity   20.7%;  Pred. No. 3.6e-07;
  Matches  488;  Conservative  190;  Mismatches  879;  Indels  804;  Gaps  113;

Qy       570 KTANSQGRRKGRITRS--MANEANSEEAITPQQSAELASMELNESSRWTEEEMETAKKGL 627

Db     16787 RTARRTPRRRRRRTRSHGSANRRRTPWALRPRRPAR---------SRWRRPPHRPGSAGA 16837

Qy       628 LEHGRNWSAIARMVGSKTVSQCKNFYFNYKKRQ-----------NLDEILQQHKLKMEK 675

Db     16838 RTRCRAATRCARFRAAGT-GRCTSPRSTWPCRRAPPAWAACPRSSGPERLPRRAPRRWRR 16896

Qy       676 ERNARRKKKKAPA------AASEEAAFPPVVEDEEMEASGVSGNEEEMVEEAEALHASGN 729

Db     16897 TRPAPRRRAPLPARSCPRRPARRRPVWPP---------------------------- 16925

Qy       730 EVPRGECSGPATVNNSSDTESIPSP---HTEAAKDTGQNGPKPPATLGADGPPPGPPTPP 786

Db     16926 --RPGRRSGPRDGRRRGAPTARPRPGRRHRSRRR--------------DGPHVGPDVPG 16968

Qy       787 RRTSRAPIEPTPASEATGAPTPPPAPP--SPS------APP----PVVPKE-----EKEE 829

Db     16969 HRWSGRPARPRPRPPGRRRPRASRSPPLRSPADVFRLRTPGTWTWPALPRNYGIGVRQSV 17028
```





RESULT 88
US-10-402-089-8
; Sequence 8, Application US/10402089
; Publication No. US20040005663A1

```
; GENERAL INFORMATION:
;  APPLICANT: Bell, Marcum  P.
;  APPLICANT:  Neff, Thomas B.
;  APPLICANT:  Polarek, James W.
;  APPLICANT:  Seeley, Todd W.
;  TITLE OF INVENTION: PORCINE COLLAGENS AND GELATINS
;  FILE REFERENCE: FP0402.3 CON
;  CURRENT APPLICATION NUMBER: US/10/402,089
;  CURRENT FILING DATE:  2003-03-26
;  PRIOR APPLICATION NUMBER: US 09/709,700
;  PRIOR FILING DATE: 2000-11-10
;  NUMBER OF SEQ ID NOS: 72
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8
;   LENGTH: 1449
;   TYPE: PRT
;   ORGANISM: Sus scrofa
US-10-402-089-8

  Query Match             2.8%;  Score 366;  DB 15;  Length 1449;
  Best Local Similarity   20.6%;  Pred. No. 1.8e-08;
  Matches  375;  Conservative  105;  Mismatches  649;  Indels  694;  Gaps   80;

Qy        703 EEMEASGVSGNEEEM-----VEEAEALH---------------ASGN------------ 729
              || :  |  | ||::      |:      |                 :||
Db         24 EEGQEEGQQGQEEDIPPVTCVQNGLRYHDRDVWKPVPCQICVCDNGNVLCDDVICDEIKN 83

Qy        730 ----EVPRGECSGPATVNNSSDTESIPSPHTEAAKDTGQNGPK----PPATLGADGPP-- 779
                  || |||         | |:   :       |||  ||:    ||   |  | |
Db         84 CPSARVPAGECCPVCPEGEVSPTDQETTGVEGPKGDTGPRGPRGPSGPPGRDGIPGQPGL 143

Qy        780 PGPPTPPRRTSRAPIEPTPASEATGAPTPP----------------------PAPPSPS 816
              |||| ||                  | | ||                      | |  ||
Db        144 PGPPGPP-----------------GPPGPPGLGGNFAPQLSYGYDEKSAGISVPGPMGPS 186

Qy        817 AP-----PPVVPKEEKEEETAAAPPVEEGE--------EQKPPAAEELAVDTGKAEEPVK 863
              |      ||  |  :  :     || | ||        : ||       | |:| :| :
Db        187 GPRGLSGPPGAPGPQGFQ----GPPGEPGEPGASGPMGPRGPPGPPGKNGDDGEAGKPGR 242

Qy        864 SECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGSGRATTAKSSGAPQDSDSSATCSA 923
                  | ||   : |     ||   |    |    |       ||    |
Db        243 -----PGERGPPGPQGARGLPGTA--GLPGMKGHRGFSGLDGAKGDAGP----------- 284

Qy        924 DEVDEAEGGDKNRLLSP----RPSLLTPTGDPRANASPQKPLDLKQLKQRAAAIPPIQVT 979
                  | |    ||     |  : | | |     |  |         | |      |
Db        285 -------AGPKGEPGSPGENGAPGQMGPRGLPGERGRPGPP-------GPAGARGNDGAT 330

Qy        980 KVHEPPREDAAPTKPA-PPAPPPPQNLQPESDAPQQPGS-SPRGKSRSPAPP-------- 1029
                 ||     || || ||  |     : |:       ||  |:|    | ||
Db        331 GAAGPP----GPTGPAGPPGPPGAVGAKGEAGPQGARGSEGPQGVRGEPGPPGPAGAAGP 386

Qy       1030 -----ADKEAFAAEAQKLPGDPPCWTSGLP-FPVPPREVIKASPHAPDPSAFSYAPPGHP 1083
                  || :     |   ||      :| | ||           |    |     :   || |
Db        387 AGNPGADGQPGGKGANGAPG-----IAGAPGFPGARGPSGPQGPSGPPGPKGNSGEPGAP 441

Qy       1084 LPLGLHDTARPVLPRPPTISNPPPLISSAKHPSVLERQIGAISQGMSVQLHVPYSEHAKA 1143
                |  ||     | |  :  ||        |:  | : |                  |:
Db        442 GSKG--DTGAKGEPGPTGVQGPP-------GPAGEEGKRG-----------------ARG 475

Qy       1144 PVGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPESLGVPTAQEASVLRGTA--LG 1201
                ||   |||                   | |: | | | | | |   :  :| |    |
Db        476 EPGPA--GLP-------------------GPPGERGGPGSRGFPGADGVAGPKGPAGERG 514

Qy       1202 SVPGGSITKGIP-STRVPSDSAITYRGSITHGTPADVLYKGTITRIIGEDSPSRLD---- 1256
                | || :  || |     | :: :     :| |:|       |       |    |:  |
Db        515 S-PGPAGPKGSPGEAGRPGEAGLPGAKGLT-GSPGSPGPDGK----TGPPGPAGQDGRPG 568

Qy       1257 -------RGREDSL----PKGHVIYEGKKGHVLSYEGGM-----SVTQCSKEDGRSSSGP 1300
                     ||:   :    |||     || |     | |:     :|    |:   . : ||
Db        569 PPGPPGARGQAGVMGFPGPKGAAGEPGKAG-----ERGVPGPPGAVGPAGKDGEAGAQGP 623

Qy       1301 PHETAAPKRTYDMMEGRVGRAISSASIEGLMGRAIPP-ERHSPHHLKEQHHIRGSITQGI 1359
                |   | |         |  | | |   :|| | | || |   |                ||:
```

```
Db        624 P-GPAGPAGE----RGEQGPA-GSPGFQGLPGPAGPPGEAGKPGE-----------QGV 665

Qy       1360 PRSYVEAQEDYLRREAKLLKREGTPPPPPPSRDLTEAYKTQALGPLKLK--PAHEGLVAT 1417
                |           | |       |   || ||:           ||      | ::|
Db        666 PGDLGAPGPSGARGERGFPGERGVQGPPGPA-----------GPRGANGAPGNDGAKG- 712

Qy       1418 VKEAGRSIHEIPREELRHTPELPLAPRPLKEGSITQGTP----LKYDTGASTTGSKKHD- 1472
                :||                 | ||     ||  || |    :  : ||:     | |
Db        713 --DAG---------------APGAP-----GS--QGAPGLQGMPGERGAAGLPGPKGDR 747

Qy       1473 ---------------VRSLIGSPGRTFPPVHPLDVMADARALERACYEESLKSRPGTAS 1516
                            || | |  |   | | |                    |   | :
Db        748 GDAGPKGADGAPGKDGVRGLTGPIGPPGPAGAPGD----------------KGETGPSG 790

Qy       1517 SSGGSIARGAPVIVPELGKPRQSPLTYED----HGAPFAGHLPRGSP---VTMREPTPRL 1569
                :| : |||||    | | |   | :       | | |   | | |    ::  | |:
Db        791 PAGPTGARGAPGDRGEPGPP--GPAGFAGPPGADGQPGAKGGPTGPPGPIGSVGAPGPKG 848

Qy       1570 QEGSLSSSKA-----SQDRKLTSTPREIAKSPHSTVPEHHPHPISPYEHLLRGVSGVDLY 1624
                 ||     |     :  |    |   | |    |        |     || :|
Db        849 ARGSAGPPGATGFPGAAGRVGPPGPSGNAGPPGPPGPAGKEGSKGP-----RGETG---- 899

Qy       1625 RSHIPLAFDPTSIPRGIPLDAAAAYYLPRHLAPNPTYPHLYPPYLIRGYPDTAALENRQT 1684
                          | | |:|         | |       |   :| |
Db        900 -------------PAGRPGEAG---------PPGPP-----GPAGEKGSPGA-------- 924

Qy       1685 IINDYITSQQMHHNTATAMAQRADMLRGLSPRESSLALNYAAGPRGIIDLSQVPHLPVLV 1744
                                  | |      :|:: :   : |    | ||
Db        925 -------------DGPAGAPGTPGPQGIAGQRGVVGLPGQRGERGF------------ 957

Qy       1745 PPTPGTPATAMDRLAYLPTAPQPFSSRHSSSPLSPGGPTHLTKPTTTSSSERERDRDRER 1804
                 || |  : :     |    |         | | ||  |  |    |  |
Db        958 ---PGLPGPSGE-----PGKQGPSGPSGERGPPGPMGPPGLAGPPGESGRE--------- 1000

Qy       1805 DRDREREKSILTSTTTVEHAPIWRPGTEQSSGSSGSSG-----GGGGSSSRPASHSHAHQ 1859
                                      || | | |  |: |     |  | : | :
Db       1001 ----------------------GAPGAEGSPGRDGAPGPKGDRGESGPAGPPGAPGAPGA 1038

Qy       1860 HSPISPRTQDALQQRPSVLHNTGMKGIITAVEPSKPTVLRSTSTSSPVRPAATFPPATHC 1919
                |: |                 | |      |: |          : || |      |
Db       1039 PGPVGP---------------AGKSGDRGETGPAGP--------AGPVGPVGARGP---- 1071

Qy       1920 PLGGTLDGVYPTLMEPVLLPKEAPRVARPERPRADTGH--------------AFLAKPP 1964
                                     | |: || | |                 : |  ||
Db       1072 -------------------------AGPQGPRGDKGETGEQGDRGIKGHRGFSGLQGPP 1105

Qy       1965 ARSGLEPASSPSKGSEPRPLVPPVSGHATIARTPAKNLAPHHASPDPPAPPASASDPHRE 2024
                 |      ||  | |                                 |  || ||  | ::
Db       1106 GPPGSPGEQGPSGASGP-------------------------AGPRGPPGSAGAPGKD 1138

Qy       2025 KTQSKPFSIQELELRSLGYHGSSYSPEGVEPVSPVSSPSLTHDKGLPKHLEELDKSHLEG 2084
                   |  |          | |          || |   |
Db       1139 GLNGLPGPI--------GPPGPRGRTGDAGPVGPPGPPG---------------------- 1169

Qy       2085 ELRPKQPGPVKLGGEAAHLPHLRPLPESQPSSSPLLQTAPGVKGHQ-------------- 2130
                   |  |||   | : :||         ||          |  | |
Db       1170 --PPGPPGPPSGGFDFSFLP--------QP---------PQEKAHDGGRYYRADDANVVR 1210

Qy       2131 --------RVVTLAQHISEVITQDYTRHHPQQLSAPLPAPLYSFPGASCPVLDLRRPPSD 2182
                      : :|:|  |  : : : :| :| :                  :|   ||:  ||
Db       1211 DRDLEVDTTLKSLSQQIENIRSPEGSRKNPAR--------------TC--RDLKMCHSD 1253

Qy       2183 LYLPPPDHGAPARGSPHSEGGKRSPEPNKTSVLGG---------GEDGIEPVSP--PEG- 2230
                              : |:   :||:   |            ||  : |  |  |:
Db       1254 W----------------KSGEYWIDPNQGCNLDAIKVFCNMETGETCVYPTQPSVPQKN 1296

Qy       2231 ---MTEPGHSRSAVYPLLYRDGEQTEPSRMGSKSPGNTSQPPAFFSKLTESNSAMVKSKK 2287
                    |   |     | | |           | | |      :||      ::
Db       1297 WYISKNPKDKRHVWYGESMTDGFQFEYG-------GEGSDPADVAIQLT-----FLRLMS 1344

Qy       2288 QEINKKLNTHNRNEPEYNISQPG 2310
                | :: :  | :|   |   | |
Db       1345 TEASQNITYHCKNSVAYMDQQTG 1367
```

```
RESULT 89
US-10-402-072A-8
; Sequence 8, Application US/10402072A
; Publication No. US20040018592A1
; GENERAL INFORMATION:
;  APPLICANT: Bell, Marcum  P.
;  APPLICANT:  Neff, Thomas B.
;  APPLICANT:  Polarek, James W.
;  APPLICANT:  Seeley, Todd W.
;  TITLE OF INVENTION: BOVINE COLLAGENS AND GELATINS
;  FILE REFERENCE: FP0402.2 CON
;  CURRENT APPLICATION NUMBER: US/10/402,072A
;  CURRENT FILING DATE:  2003-03-26
;  PRIOR APPLICATION NUMBER: US 09/709,700
;  PRIOR FILING DATE: 2000-11-10
;  NUMBER OF SEQ ID NOS: 72
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8
;   LENGTH: 1449
;   TYPE: PRT
;   ORGANISM: Sus scrofa
US-10-402-072A-8

  Query Match             2.8%;  Score 366;  DB 15;  Length 1449;
  Best Local Similarity   20.6%;  Pred. No. 1.8e-08;
  Matches  375;  Conservative  105;  Mismatches  649;  Indels  694;  Gaps   80;

Qy        703 EEMEASGVSGNEEEM-----VEEAEALH---------------ASGN------------- 729
Db         24 EEGQEEGQQGQEEDIPPVTCVQNGLRYHDRDVWKPVPCQICVCDNGNVLCDDVICDEIKN 83

Qy        730 ----EVPRGECSGPATVNNSSDTESIPSPHTEAAKDTGQNGPK----PPATLGADGPP-- 779
Db         84 CPSARVPAGECCPVCPEGEVSPTDQETTGVEGPKGDTGPRGPRGPSGPPGRDGIPGQPGL 143

Qy        780 PGPPTPPRRTSRAPIEPTPASEATGAPTPP-----------------------PAPPSPS 816
Db        144 PGPPGPP-----------------GPPGPPGLGGNFAPQLSYGYDEKSAGISVPGPMGPS 186

Qy        817 AP-----PPVVPKEEKEEETAAAPPVEEGE--------EQKPPAAEELAVDTGKAEEPVK 863
Db        187 GPRGLSGPPGAPGPQGFQ----GPPGEPGEPGASGPMGPRGPPGPPGKNGDDGEAGKPGR 242

Qy        864 SECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGSGRATTAKSSGAPQDSDSSATCSA 923
Db        243 -----PGERGPPGPQGARGLPGTA--GLPGMKGHRGFSGLDGAKGDAGP----------- 284

Qy        924 DEVDEAEGGDKNRLLSP----RPSLLTPTGDPRANASPQKPLDLKQLKQRAAAIPPIQVT 979
Db        285 -------AGPKGEPGSPGENGAPGQMGPRGLPGERGRPGPP-------GPAGARGNDGAT 330

Qy        980 KVHEPPREDAAPTKPA-PPAPPPPQNLQPESDAPQQPGS-SPRGKSRSPAPP-------- 1029
Db        331 GAAGPP----GPTGPAGPPGFPGAVGAKGEAGPQGARGSEGPQGVRGEPGPPGPAGAAGP 386

Qy       1030 -----ADKEAFAAEAQKLPGDPPCWTSGLP-FPVPPREVIKASPHAPDPSAFSYAPPGHP 1083
Db        387 AGNPGADGQPGGKGANGAPG-----IAGAPGFPGARGPSGPQGPSGPPGPKGNSGEPGAP 441

Qy       1084 LPLGLHDTARPVLPRPPTISNPPPLISSAKHPSVLERQIGAISQGMSVQLHVPYSEHAKA 1143
Db        442 GSKG--DTGAKGEPGPTGVQGPP-------GPAGEEGKRG-----------------ARG 475

Qy       1144 PVGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPESLGVPTAQEASVLRGTA--LG 1201
Db        476 EPGPA--GLP------------------GPPGERGGPGSRGFPGADGVAGPKGPAGERG 514

Qy       1202 SVPGGSITKGIP-STRVPSDSAITYRGSITHGTPADVLYKGTITRIIGEDSPSRLD---- 1256
Db        515 S-PGPAGPKGSPGEAGRPGEAGLPGAKGLT-GSPGSPGPDGK----TGPPGPAGQDGRPG 568

Qy       1257 -------RGREDSL----PKGHVIYEGKKGHVLSYEGGM-----SVTQCSKEDGRSSSGP 1300
Db        569 PPGPPGARGQAGVMGFPGPKGAAGEPGKAG-----ERGVPGPPGAVGPAGKDGEAGAQGP 623

Qy       1301 PHETAAPKRTYDMMEGRVGRAISSASIEGLMGRAIPP-ERHSPHHLKEQHHIRGSITQGI 1359
Db        624 P-GPAGPAGE----RGEQGPA-GSPGFQGLPGPAGPPGEAGKPGE-----------QGV 665

Qy       1360 PRSYVEAQEDYLRREAKLLKREGTPPPPPPSRDLTEAYKTQALGPLKLK--PAHEGLVAT 1417
Db        666 PGDLGAPGPSGARGERGFPGERGVQGPPGPA-----------GPRGANGAPGNDGAKG- 712

Qy       1418 VKEAGRSIHEIPREELRHTPELPLAPRPLKEGSITQGTP----LKYDTGASTTGSKKHD- 1472
Db        713 --DAG---------------APGAP-----GS--QGAPGLQGMPGERGAAGLPGPKGDR 747

Qy       1473 ----------------VRSLIGSPGRTFPPVHPLDVMADARALERACYEESLKSRPGTAS 1516
Db        748 GDAGPKGADGAPGKDGVRGLTGPIGPPGPAGAPGD----------------KGETGPSG 790

Qy       1517 SSGGSIARGAPVIVPELGKPRQSPLTYED----HGAPFAGHLPRGSP---VTMREPTPRL 1569
Db        791 PAGPTGARGAPGDRGEPGPP--GPAGFAGPPGADGQPGAKGGPTGPPGPIGSVGAPGPKG 848

Qy       1570 QEGSLSSSKA-----SQDRKLTSTPREIAKSPHSTVPEHHPHPISPYEHLLRGVSGVDLY 1624
Db        849 ARGSAGPPGATGFPGAAGRVGPPGPSGNAGPPGPPGPAGKEGSKGP-----RGETG---- 899

Qy       1625 RSHIPLAFDPTSIPRGIPLDAAAAYYLPRHLAPNPTYPHLYPPYLIRGYPDTAALENRQT 1684
Db        900 -------------PAGRPGEAG---------PPGPP-----GPAGEKGSPGA-------- 924

Qy       1685 IINDYITSQQMHHNTATAMAQRADMLRGLSPRESSLALNYAAGPRGIIDLSQVPHLPVLV 1744
Db        925 --------------DGPAGAPGTPGPQGIAGQRGVVGLPGQRGERGF------------- 957

Qy       1745 PPTPGTPATAMDRLAYLPTAPQPFSSRHSSSPLSPGGPTHLTKPTTTSSSERERDRDRER 1804
Db        958 ---PGLPGPSGE-----PGKQGPSGPSGERGPPGPMGPPGLAGPPGESGRE--------- 1000

Qy       1805 DRDREREKSILTSTTTVEHAPIWRPGTEQSSGSSGSSG-----GGGGSSSRPASHSHAHQ 1859
Db       1001 ---------------------GAPGAEGSPGRDGAPGPKGDRGESGPAGPPGAPGAPGA 1038

Qy       1860 HSPISPRTQDALQQRPSVLHNTGMKGIITAVEPSKPTVLRSTSTSSPVRPAATFPPATHC 1919
Db       1039 PGPVGP---------------AGKSGDRGETGPAGP--------AGPVGPVGARGP---- 1071

Qy       1920 PLGGTLDGVYPTLMEPVLLPKEAPRVARPERPRADTGH---------------AFLAKPP 1964
Db       1072 ---------------------------AGPQGPRGDKGETGEQGDRGIKGHRGFSGLQGPP 1105

Qy       1965 ARSGLEPASSPSKGSEPRPLVPPVSGHATIARTPAKNLAPHHASPDPPAPPASASDPHRE 2024
Db       1106 GPPGSPGEQGPSGASGP---------------------------AGPRGPPGSAGAPGKD 1138

Qy       2025 KTQSKPFSIQELELRSLGYHGSSYSPEGVEPVSPVSSPSLTHDKGLPKHLEELDKSHLEG 2084
Db       1139 GLNGLPGPI--------GPPGPRGRTGDAGPVGPPGPPG--------------------- 1169

Qy       2085 ELRPKQPGPVKLGGEAAHLPHLRPLPESQPSSSPLLQTAPGVKGHQ-------------- 2130
Db       1170 --PPGPPGPPSGGFDFSFLP--------QP---------PQEKAHDGGRYYRADDANVVR 1210

Qy       2131 --------RVVTLAQHISEVITQDYTRHHPQQLSAPLPAPLYSFPGASCPVLDLRRPPSD 2182
Db       1211 DRDLEVDTTLKSLSQQIENIRSPEGSRKNPAR---------------TC--RDLKMCHSD 1253

Qy       2183 LYLPPPDHGAPARGSPHSEGGKRSPEPNKTSVLGG---------GEDGIEPVSP--PEG- 2230
Db       1254 W----------------KSGEYWIDPNQGCNLDAIKVFCNMETGETCVYPTQPSVPQKN 1296

Qy       2231 ---MTEPGHSRSAVYPLLYRDGEQTEPSRMGSKSPGNTSQPPAFFSKLTESNSAMVKSKK 2287
```

```
Db       1297 WYISKNPKDKRHVWYGESMTDGFQFEYG-------GEGSDPADVAIQLT-----FLRLMS 1344

Qy       2288 QEINKKLNTHNRNEPEYNISQPG 2310

Db       1345 TEASQNITYHCKNSVAYMDQQTG 1367


RESULT 90
US-10-342-331-5
; Sequence 5, Application US/10342331
; Publication No. US20030229205A1
; GENERAL INFORMATION:
;  APPLICANT: VAN HEERDE, GEORGE V.
;  APPLICANT:  VAN RIJN, ALEXIS C.
;  APPLICANT:  BOUWSTRA, JAN B.
;  APPLICANT:  DE WOLF, FREDERIK A.
;  APPLICANT:  MOOBROEK, ANDREAS
;  APPLICANT:  WERTEN, MARC W.T.
;  APPLICANT:  WIND, RICHELE D.
;  APPLICANT:  VAN DEN BOSCH, TANJA J.
;  TITLE OF INVENTION: SILVER HALIDE EMULSIONS WITH RECOMBINANT COLLAGEN
;  TITLE OF INVENTION:  SUITABLE FOR PHOTOGRAPHIC APPLICATION AND ALSO THE
;  TITLE OF INVENTION:  PREPARATION THEREOF
;  FILE REFERENCE: 2728-2
;  CURRENT APPLICATION NUMBER: US/10/342,331
;  CURRENT FILING DATE:  2003-01-15
;  PRIOR APPLICATION NUMBER: US/09/219,849
;  PRIOR FILING DATE: 1998-12-23
;  NUMBER OF SEQ ID NOS: 50
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 960
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Illustrative
;   OTHER INFORMATION: amino acid sequence
US-10-342-331-5

  Query Match             2.8%;  Score 365;  DB 15;  Length 960;
  Best Local Similarity   21.3%;  Pred. No. 1.3e-08;
  Matches  337;  Conservative   33;  Mismatches  490;  Indels  720;  Gaps   78;

Qy        726 ASGNEVPRGECSGPATVNNSSD--TESIPSPHTEAAKDTGQNGPKPPATLGADGPP---- 779

Db         11 AHGPAGPKG-APGPAGPPGSRDPGPPGAPGP----AGPPGSRDPGPPGAPGPAGPPGSRD 65

Qy        780 PGPP-------TPPRRTSRAPIEPTPASE-------ATGAPTP--PPA---PPSPSAPPP 820

Db         66 PGPPGAPGPAGPPGSRDPGPPGAPGPAGPPGSRDPGPPGAPGPAGPPGSRDPGPPGAPGP 125

Qy        821 VVPKEEKEEETAAAPPVEEGEEQKPPAAEELAVDTGKAEEPVKSECTEEAEEGPAKGKDA 880

Db        126 AGPPGSRDPGPPGAP---------GPAGPPGSRDPGPPGAP--------GPAGPPGSRDP 168

Qy        881 EAAEATAEGALKAEKKEGGSGRATTAKSSGAPQDSDSSATCSADEVDEAEGGDKNRLLSP 940

Db        169 GPPGAPGPAGPPGSRDPGPPGAPGPAGPPGSRDPGPPGAPGPAGPPGSRDPGP------- 221

Qy        941 RPSLLTPTGDPRANASPQKPLDLKQLKQRAAAIPPIQVTKVHEPPREDAAP--------T 992

Db        222 -PGAHGPAG-PKGAHGPAGP---------KGAHGPAGPKGAHGPAGPKGAPGPAGPPGSR 270

Qy        993 KPAPPAPPPPQNLQPESDAPQQPGS-SPRGK--SRSPAPPADKEAFAAEAQKLPGDPPCW 1049

Db        271 DPGPPGAPGPAG-PPGSRDPGPPGAPGPAGPPGSRDPGPPGAPGPAGPPGSRDPGPP--- 326

Qy       1050 TSGLPFPVPPREVIKASPHAPDPSAFSYAPPGHPLPLGLHDTARPVLPRPPTISNPPPLI 1109

Db        327 --GAPGPAGP-------PGSRDP-----GPPGAPGPAGPPGSRDPGPPGAPGPAGPP--- 369

Qy       1110 SSAKHPSVLERQIGAISQGMSVQLHVPYSEHAKAPVGPVTMGLPLPMDPKKLAPFSGVKQ 1169

Db        370 -----------GSRDPG---------PPGAPGPAGP-----PGSRDP---------- 391

Qy       1170 EQLSPRGQAGPPESLGVPTAQEASVLRGTALGSVPGGSITKGIPSTRVPSDSAITYRGSI 1229

Db        392 -----GPPGAPGPAGPPGSRDP---------GPPGAPGPAGPPGSRDPGPP-------- 428

Qy       1230 THGTPADVLYKGTITRIIGEDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQC 1289

Db        429 --GAP-----------GPAGPP----GSRDPGPPG---APGPAG-------------- 452

Qy       1290 SKEDGRSSSGPP--HETAAPKRTYDMMEGRVGRAISSASIEGLMGRAIPPERHSPHHLKE 1347

Db        453 --PPGSRDPGPPGAHGPAGPK-------------------GAHGPAGPKGAHGPAGPKG 490

Qy       1348 QHHIRGSITQGIPRSYVEAQEDYLRREAKLLKREGTPPP--PPPSRDLTEAYKTQALGPL 1405

Db        491 AHGPAGP-------------------------KGAPGPAGPPGSRD------------- 511

Qy       1406 KLKPAHEGLVATVKEAGRSIHEIPREELRHTPELPLAPRPLKEGSITQGTPLKYDTGAST 1465

Db        512 ------------------------------PGPPGAPGP-------AGPPGSRDPGPP- 532

Qy       1466 TGSKKHDVRSLIGSPGRTFPPVHPLDVMADARALERACYEESLKSRPGTASSSGGSIARG 1525

Db        533 -----------GAPGPAGPP--------------------------------GSRDPG 547

Qy       1526 APVIVPELGKPRQSPLTYEDHGAPFAGHLPRGSPVTMREPTPRLQEGSLSSSKASQDRKL 1585

Db        548 PPGAPGPAGPPGS-----RDPGPPGAPG-PAGPPGS-RDPGPPGAPGP-AGPPGSRD--- 596

Qy       1586 TSTPREIAKSPHSTVPEHHPHPISPYEHLLRGVSGVDLYRSHIPLAFDPTSIPRGIPLDA 1645

Db        597 -------------PGPPGAPGPAGP-----------------PGSRDPG--PPG----- 618

Qy       1646 AAAYYLPRHLAPNPTYPHLYPPYLIRGYPDTAALENRQTIINDYITSQQMHHNTATAMAQ 1705

Db        619 ----------APGP---------------------------------------------- 622

Qy       1706 RADMLRGLSPRESSLALNYAAGPRGIIDLSQVPHLPVLVPPTPGTPATA---MDRLAYLP 1762

Db        623 --------------------AGPPGSRD----------PGPPGAPGPAGPPGSRDPGPP 651

Qy       1763 TAPQPFSSRHSSSPLSPGGPTHLTKPTTTSSSERERDRDRERDRDREREKSILTSTTTVE 1822

Db        652 GAPGPAGPPGSRDPGPPGAP-------------------------------------- 671

Qy       1823 HAPIWRPGTEQSSGSSGSSGGGGGSSSR----PASHSHA---HQHSPISPRTQDALQQRP 1875

Db        672 -GPAGPPGS-RDPGPPGAPGPAGPPGSRDPGPPGAHGPAGPKGAHGPAGPK--------- 720

Qy       1876 SVLHNTGMKGIITAVEPSKPTVLRSTSTSSPVRPAATFPPATHCPLGGTLDGVYPTLMEP 1935

Db        721 ------GAHG----------------------PAG--PKGAHGPAG------------- 736

Qy       1936 VLLPKEAPRVARPERPRADTGHAFLAKPPARSGLEPASSPSKGSE-------PRPLVPPV 1988

Db        737 ---PKGAPGPAGPPGSR-DPG------PPGAPG--PAGPP--GSRDPGPPGAPGPAGPPG 782

Qy       1989 SGHATIARTPAKNLAPHHASP----DP-------PAPPASASDPHREKTQSKPFSIQELE 2037

Db        783 S------RDPGPPGAPGPAGPPGSRDPGPPGAPGPAGPPGSRDP---------------- 820

Qy       2038 LRSLGYHGSSYSPEGVE-PVSPVSSPSLTHDKGLPKHLEELDKSHLEGELRPKQPGPVKL 2096

Db        821 -----------GPPGAPGPAGPPGS----RDPG-----------------PPGAPGPAGP 848

Qy       2097 GGEAAHLPHLRPLPESQPSS-SPLLQTAPGVKGHQRVVTLAQHISEVITQDYTRHHPQQL 2155

Db        849 PGSRDPGPPGAPGPAGPPGSRDPGPPGAPGPAG----------------------PPGS 885

Qy       2156 SAPLPAPLYSFPGASCPVLDLRRPPSDLYLPPPDHGAPARGSPHSEGGKRSPEPNKTSVL 2215
```

```
Db        886 RDPGP------PGAPGPA----GPPGSRDPGPP--GAPGPAGP---PGSRDPGP------ 924

Qy       2216 GGGEDGIEPVSPPEGMTEPG 2235
Db        925 -PGAPG--PAGPP-GSRDPG 940
```

RESULT 91
US-10-664-859-15
; Sequence 15, Application US/10664859
; Publication No. US20040038901A1
; GENERAL INFORMATION:
; APPLICANT: BASLER, Konrad
; APPLICANT: BRUNNER, Erich
; APPLICANT: FROESCH, Barbara
; APPLICANT: KRAMPS, Thomas
; APPLICANT: PETER, Oliver
; TITLE OF INVENTION: ESSENTIAL DOWNSTREAM COMPONENT OF THE WINGLESS SIGNALING PATHWAY
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC APPLICATIONS BASED THEREON
; FILE REFERENCE: Q60361
; CURRENT APPLICATION NUMBER: US/10/664,859
; CURRENT FILING DATE: 2003-09-22
; PRIOR APPLICATION NUMBER: US/09/915,543
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 60/221,502
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 1426
; TYPE: PRT
; ORGANISM: Human lgs/bcl9
US-10-664-859-15

Query Match 2.8%; Score 364; DB 12; Length 1426;
Best Local Similarity 19.7%; Pred. No. 2.2e-08;
Matches 377; Conservative 197; Mismatches 685; Indels 658; Gaps 98;

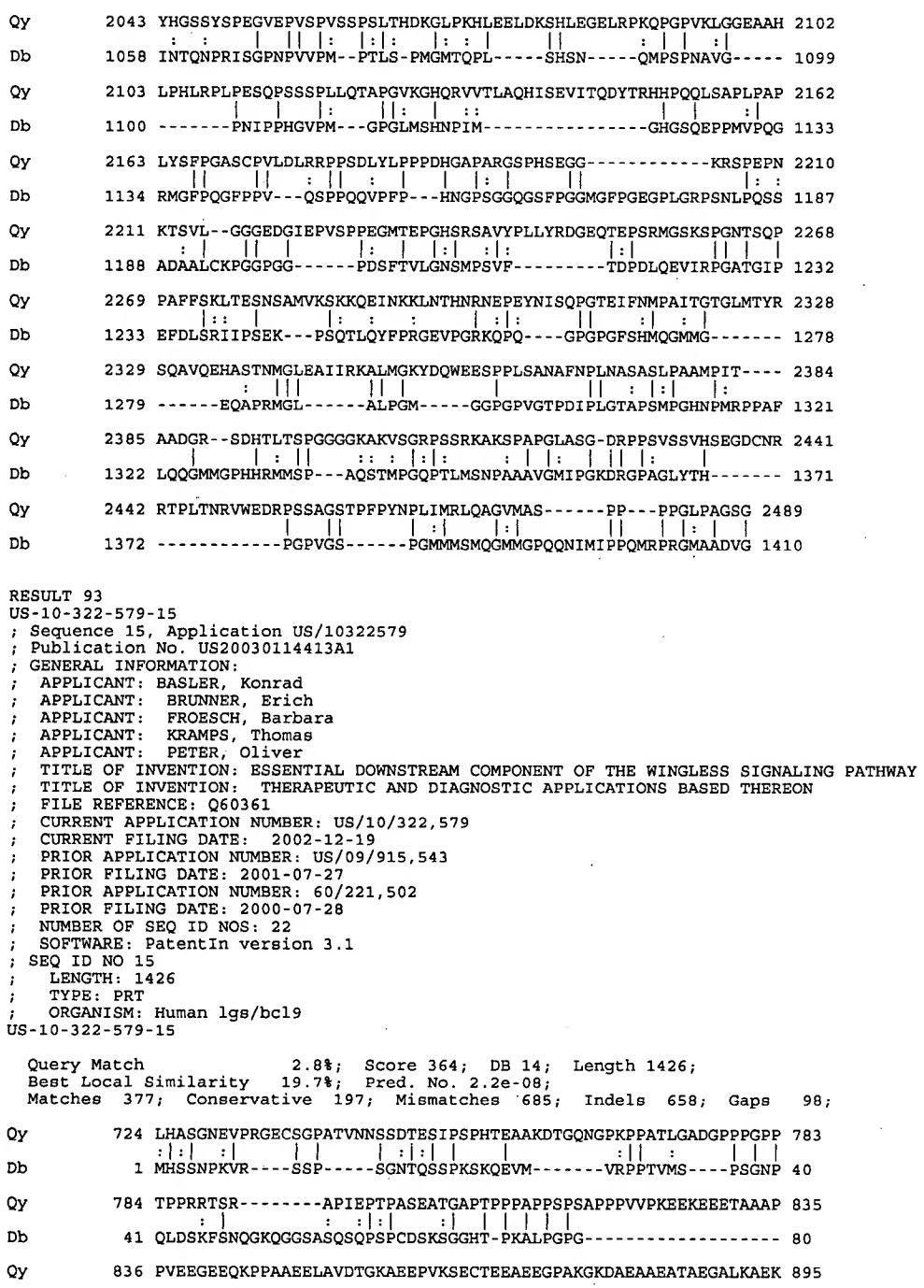
```
Qy        724 LHASGNEVPRGECSGPATVNNSSDTESIPSPHTEAAKDTGQNGPKPPATLGADGPPPGPP 783
Db          1 MHSSNPKVR----SSP-----SGNTQSSPKSKQEVM-------VRPPTVMS----PSGNP 40

Qy        784 TPPRRTSR--------APIEPTPASEATGAPTPPPAPPSPSAPPPVVPKEEKEEETAAAP 835
Db         41 QLDSKFSNQGKQGGSASQSQPSPCDSKSGGHT-PKALPGPG------------------- 80

Qy        836 PVEEGEEQKPPAAEELAVDTGKAEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEK 895
Db         81 ----GSMGLKNGAGNGAKGKGKRERSISADSFDQRDPG-TPNDDSDIKECNSADHIKSQD 135

Qy        896 KEGGSGRAT----TAKSSGAPQDSDSSATCSADEVDEAEGGDKNRLLSPRPSLLTPT--- 948
Db        136 SQHTPHSMTPSNATAPRSSTPSHGQTTAT-----------------EPTPAQKTPAKVV 177

Qy        949 ---GDPRANASPQKPL-------------DLKQLK-QRAAAIPPIQVTKVHEPPREDAAP 991
Db        178 YVFSTEMANKAAEAVLKGQVETIVSFHIQNISNNKTERSTAPLNTQISALRNDP------ 231

Qy        992 TKPAPPAPPPPQNLQPESD-----APQQPGSSPRGKSRSPAPPADKEAFAAEAQKLP--G 1044
Db        232 -KPLPQQPPAPANQDQNSSQNTRLQPTPPIPAPAPKPAAPPRPLDRESPGVENKLIPSVG 290

Qy       1045 DPPCWTSGLPFPVPP-----------REVIKASPHAPDPSAFSYAPPGHPLPLGLHDTAR 1093
Db        291 SPASST-----PLPPDGTGPNSTPNNRAVTPVSQGSNSSSADPKAPPPPPVSSG------ 339

Qy       1094 PVLPRPPTISNPPPLISSAKHPSVLERQIGAISQGMSVQLHVPYSEHAKAPVGPVTMGLP 1153
Db        340 ----EPPTLGENPDGLSQEQ----LEHRERSLQTLRDIQ---------------RMLFP 375

Qy       1154 LPMDPKKLAPFSGVKQEQLSPRGQAGPPESLGV---PTAQ-----EASVLRGTALGSVPG 1205
Db        376 ---DEKE---FTGAQ--------SGGPQQNPGVLDGPQKKPEGPIQAMMAQSQSLGKGPG 421

Qy       1206 GSITKGI-------------PSTRVP----SDSAITYRGSITHGTPADVLYKGTITRIIG 1248
Db        422 PRTDVGAPFGPQGHRDVPFSPDEMVPPSMNSQSGTIGPDHLDHMTPEQIAW--------- 472

Qy       1249 EDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGR-SSSGPPHETAAP 1307
Db        473 --------------LKLQQEFYEEKR----RKQEQVVVQQCSLQDMMVHQHGPRGVVRGP 514

Qy       1308 KRTYDM--MEGRVGRAISSASIEGL-MGRAIPPERHSPHHLKEQHHIRGSITQGIPRSYV 1364
Db        515 PPPYQMTPSEGWAPGGTEPFS-DGINMPHSLPPRGMAPH---------------------- 552

Qy       1365 EAQEDYLRREAKLLKREGTPPPPPPSRDLTEAYKTQALGPLKLKPAHEGLVATVKEAGRS 1424
Db        553 --------------------PNMPGSQ-------------MRL-PGFAGMINSEMEG--- 575

Qy       1425 IHEIPREELRHTPELPLAPRPLKEGSITQGTPLKYDTGASTTGSKKHDVRSLIGSPGR-- 1482
Db        576 ---------------PNVPNPASRPGLS-GVSWPDDVPKIPDGRNFPPGQGIFSGPGRGE 619

Qy       1483 TFPPVHPLDVMADARALERACYEESLKSR-----PGTASSSGGSIARGAPVIVPELGKPR 1537
Db        620 RFP---------NPQGLSEEMFQQQLAEKQLGLPPGMAMEG----------IRPSMEMNR 660

Qy       1538 QSPLTYEDHGAPFAGHL-PRGSPVTMREPTPRLQEGSLSSSKASQDRKLTSTPREIAKSP 1596
Db        661 MIPGSQR--------HMEPGNNPIFPRIPV----EGPLSPSRGD---------------- 692

Qy       1597 HSTVPEHHPHPISPYEHLLRGV--SGVDLYRSHIPLAFDPTSIPRGIP---LDAAAAYYL 1651
Db        693 ---FPKGIPPQMGPGRELEFGMVPSGM---KGDVNLNVNMGSNSQMIPQKMREAGAG--- 743

Qy       1652 PRHLAPNPTYPHLYPPYLIRGYPDTAALENRQTIINDYITSQQMHHNTATAMAQRADMLR 1711
Db        744 --------------PEEMLKLRPGG----------SDMLPAQQ----------KMVPLPF 769

Qy       1712 GLSPRESSLALNYAAGPRGIIDLSQVP----HLPVLVPPTPGTPATAMDRLAYLPTAPQP 1767
Db        770 GEHPQQ-----EYGMGPRPFLPMSQGPGSNSGLRNLREPI-GPDQRTNSRLSHMPPLPLN 823

Qy       1768 FSSRHSSSPLSPGGPTHLTKPTTTSSSERERDRDRERDRDRDRREREKSILTSTTTVEHAPIW 1827
Db        824 PSSNPTSLNTAP--------PVQRGLGRKPLD-------------------ISVAGSQVH 856

Qy       1828 RPGTEQSSGSSGSSGGGGGSSSRPASHSHAHQ-HSPI--SPRTQDALQQRPSVLHNTGM- 1883
Db        857 SPGI-----------------NPLKSPTMHQVQSPMLGSPSGNLKSPQTPSQL--AGML 896

Qy       1884 KGIITAVEPSKPTVLRSTSTSSPVRPAATFPPATHCPLGGTLDGVYPTLMEPVLLPK-EA 1942
Db        897 AGPAAAASIKSPPVLGSAAAS----PVHLKSPSLPAPSPGWTSSPKPPLQSPGIPPNHKA 952

Qy       1943 P-RVARPERPRADTGHAFLAKPPARSGLEPASSPSKGSEPRPLVPPVSGHATIARTP--- 1998
Db        953 PLTMASP----AMLGNVESGGPPPPTASQPASVNIPGSLPSSTPYTMPPEPTLSQNPLSI 1008

Qy       1999 -----AKNLAP-----HH------ASPDPPAPPASASDPHREKTQSKPFSIQELELRSLG 2042
Db        1009 MMSRMSKFAMPSSTPLYHDAIKTVASSDDDSPPARS--PNLPSMNNMP---------GMG 1057

Qy       2043 YHGSSYSPEGVEPVSPVSSPSLTHDKGLPKHLEELDKSHLEGELRPKQPGPVKLGGEAAH 2102
Db        1058 INTQNPRISGPNPVVPM--PTLS-PMGMTQPL-----SHSN-----QMPSPNAVG----- 1099

Qy       2103 LPHLRPLPESQPSSSPLLQTAPGVKGHQRVVTLAQHISEVITQDYTRHHPQQLSAPLPAP 2162
Db        1100 -------PNIPPHGVPM---GPGLMSHNPIM---------------GHGSQEPPMVPQG 1133

Qy       2163 LYSFPGASCPVLDLRRPPSDLYLPPPDHGAPARGSPHSEGG------------KRSPEPN 2210
Db        1134 RMGFPQGFPPV---QSPPQQVPFP---HNGPSGGQGSFPGGMGFPGEGPLGRPSNLPQSS 1187
```

```
Qy       2211 KTSVL--GGGEDGIEPVSPPEGMTEPGHSRSAVYPLLYRDGEQTEPSRMGSKSPGNTSQP 2268
Db       1188 ADAALCKPGGPGG------PDSFTVLGNSMPSVF---------TDPDLQEVIRPGATGIP 1232

Qy       2269 PAFFSKLTESNSAMVKSKKQEINKKLNTHNRNEPEYNISQPGTEIFNMPAITGTGLMTYR 2328
Db       1233 EFDLSRIIPSEK---PSQTLQYFPRGEVPGRKQPQ----GPGPGFSHMQGMMG------- 1278

Qy       2329 SQAVQEHASTNMGLEAIIRKALMGKYDQWEESPPLSANAFNPLNASASLPAAMPIT---- 2384
Db       1279 ------EQAPRMGL------ALPGM-----GGPGPVGTPDIPLGTAPSMPGHNPMRPPAF 1321

Qy       2385 AADGR--SDHTLTSPGGGGKAKVSGRPSSRKAKSPAPGLASG-DRPPSVSSVHSEGDCNR 2441
Db       1322 LQQGMMGPHHRMMSP---AQSTMPGQPTLMSNPAAAVGMIPGKDRGPAGLYTH------- 1371

Qy       2442 RTPLTNRVWEDRPSSAGSTPFPYNPLIMRLQAGVMAS------PP---PPGLPAGSG 2489
Db       1372 ------------PGPVGS------PGMMMSMQGMMGPQQNIMIPPQMRPRGMAADVG 1410
```

RESULT 92
US-09-915-543-15
; Sequence 15, Application US/09915543
; Publication No. US20020086986A1
; GENERAL INFORMATION:
; APPLICANT: BASLER, Konrad
; APPLICANT: BRUNNER, Erich
; APPLICANT: FROESCH, Barbara
; APPLICANT: KRAMPS, Thomas
; APPLICANT: PETER, Oliver
; TITLE OF INVENTION: ESSENTIAL DOWNSTREAM COMPONENT OF THE WINGLESS SIGNALING PATHWAY
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC APPLICATIONS BASED THEREON
; FILE REFERENCE: Q60361
; CURRENT APPLICATION NUMBER: US/09/915,543
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 60/221,502
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 1426
; TYPE: PRT
; ORGANISM: Human lgs/bcl9
US-09-915-543-15

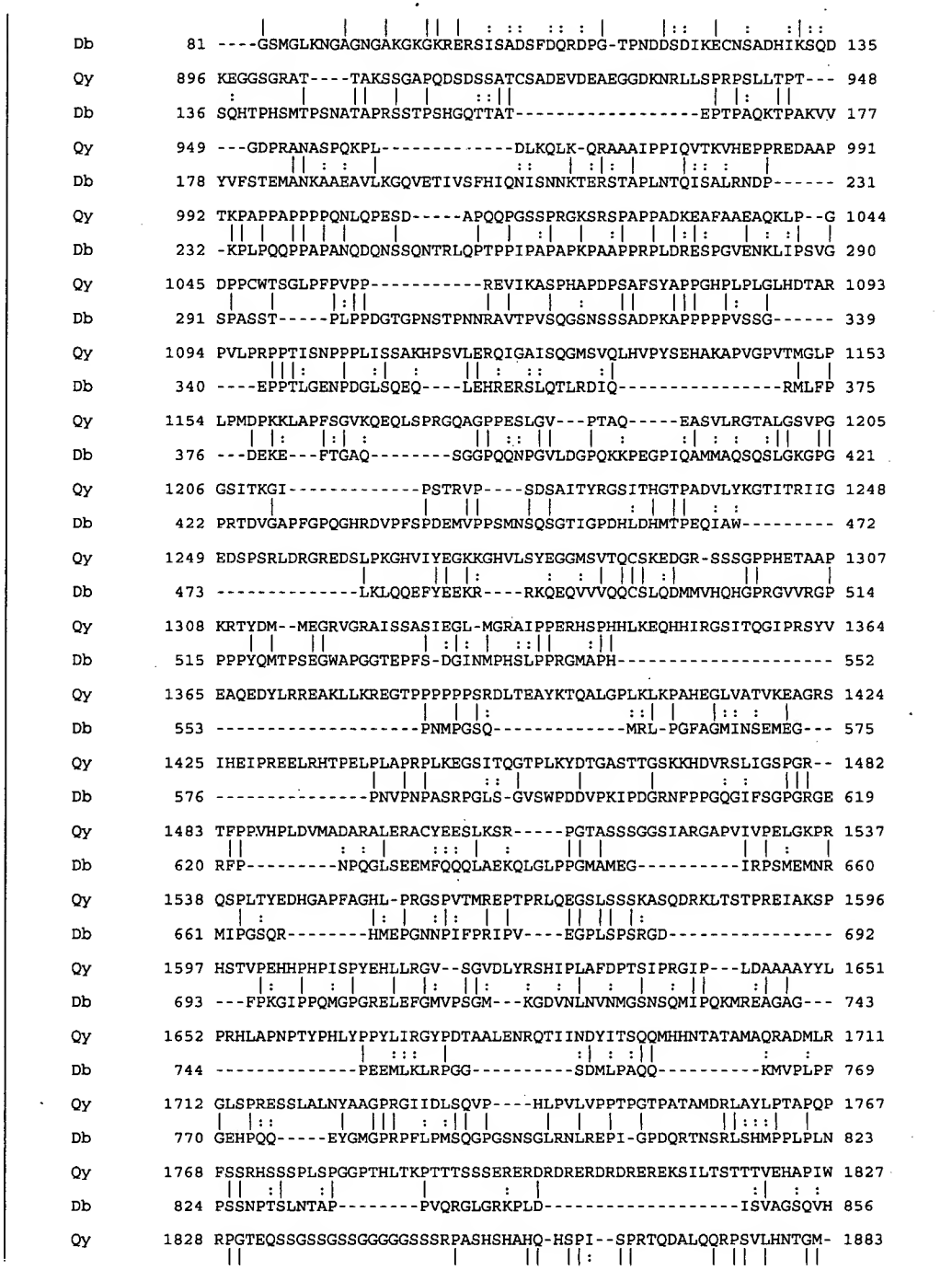
Query Match 2.8%; Score 364; DB 12; Length 1426;
Best Local Similarity 19.7%; Pred. No. 2.2e-08;
Matches 377; Conservative 197; Mismatches 685; Indels 658; Gaps 98;

```
Qy        724 LHASGNEVPRGECSGPATVNNSSDTESIPSPHTEAAKDTGQNGPKPPATLGADGPPPGPP 783
Db          1 MHSSNPKVR----SSP-----SGNTQSSPKSKQEVM-------VRPPTVMS----PSGNP 40

Qy        784 TPPRRTSR--------APIEPTPASEATGAPTPPPAPPSPSAPPPVVPKEEKEEETAAAP 835
Db         41 QLDSKFSNQGKQGGSASQSQPSPCDSKSGGHT-PKALPGPG------------------ 80

Qy        836 PVEEGEEQKPPAAEELAVDTGKAEEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEK 895
Db         81 ----GSMGLKNGAGNGAKGKGKRERSISADSFDQRDPG-TPNDDSDIKECNSADHIKSQD 135

Qy        896 KEGGSGRAT----TAKSSGAPQDSDSSATCSADEVDEAEGGDKNRLLSPRPSLLTPT--- 948
Db        136 SQHTPHSMTPSNATAPRSSTPSHGQTTAT-----------------EPTPAQKTPAKVV 177

Qy        949 ---GDPRANASPQKPL------------DLKQLK-QRAAAIPPIQVTKVHEPPREDAAP 991
Db        178 YVFSTEMANKAAEAVLKGQVETIVSFHIQNISNNKTERSTAPLNTQISALRNDP------ 231

Qy        992 TKPAPPAPPPPQNLQPESD-----APQQPGSSPRGKSRSPAPPADKEAFAAEAQKLP--G 1044
Db        232 -KPLPQQPPAPANQDQNSSQNTRLQPTPPIPAPAPKPAAPPRPLDRESPGVENKLIPSVG 290

Qy       1045 DPPCWTSGLPFPVPP----------REVIKASPHAPDPSAFSYAPPGHPLPLGLHDTAR 1093
Db        291 SPASST-----PLPPDGTGPNSTPNNRAVTPVSQGSNSSSADPKAPPPPPVSSG------ 339

Qy       1094 PVLPRPPTISNPPPLISSAKHPSVLERQIGAISQGMSVQLHVPYSEHAKAPVGPVTMGLP 1153
Db        340 ----EPPTLGENPDGLSQEQ----LEHRERSLQTLRDIQ---------------RMLFP 375

Qy       1154 LPMDPKKLAPFSGVKQEQLSPRGQAGPPESLGV---PTAQ-----EASVLRGTALGSVPG 1205
Db        376 ---DEKE---FTGAQ--------SGGPQQNPGVLDGPQKKPEGPIQAMMAQSQSLGKGPG 421

Qy       1206 GSITKGI-------------PSTRVP----SDSAITYRGSITHGTPADVLYKGTITRIIG 1248
Db        422 PRTDVGAPFGPQGHRDVPFSPDEMVPPSMNSQSGTIGPDHLDHMTPEQIAW--------- 472

Qy       1249 EDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKEDGR-SSSGPPHETAAP 1307
Db        473 -------------LKLQQEFYEEKR----RKQEQVVVQQCSLQDMMVHQHGPRGVVRGP 514

Qy       1308 KRTYDM--MEGRVGRAISSASIEGL-MGRAIPPERHSPHHLKEQHHIRGSITQGIPRSYV 1364
Db        515 PPPYQMTPSEGWAPGGTEPFS-DGINMPHSLPPRGMAPH---------------------- 552

Qy       1365 EAQEDYLRREAKLLKREGTPPPPPPSRDLTEAYKTQALGPLKLKPAHEGLVATVKEAGRS 1424
Db        553 --------------------PNMPGSQ------------MRL-PGFAGMINSEMEG--- 575

Qy       1425 IHEIPREELRHTPELPLAPRPLKEGSITQGTPLKYDTGASTTGSKKHDVRSLIGSPGR-- 1482
Db        576 --------------PNVPNPASRPGLS-GVSWPDDVPKIPDGRNFPPGQGIFSGPGRGE 619

Qy       1483 TPPPVHPLDVMADARALERACYEESLKSR-----PGTASSSGGSIARGAPVIVPELGKPR 1537
Db        620 RFP---------NPQGLSEEMFQQQLAEKQLGLPPGMAMEG----------IRPSMEMNR 660

Qy       1538 QSPLTYEDHGAPFAGHL-PRGSPVTMREPTPRLQEGSLSSSKASQDRKLTSTPREIAKSP 1596
Db        661 MIPGSQR--------HMEPGNNPIFPRIPV----EGPLSPSRGD--------------- 692

Qy       1597 HSTVPEHHPHPISPYEHLLRGV--SGVDLYRSHIPLAFDPTSIPRGIP---LDAAAAYYL 1651
Db        693 ---FPKGIPPQMGPGRELEFGMVPSGM---KGDVNLNVNMGSNSQMIPQKMREAGAG--- 743

Qy       1652 PRHLAPNPTYPHLYPPYLIRGYPDTAALENRQTIINDYITSQQMHHNTATAMAQRADMLR 1711
Db        744 -------------PEEMLKLRPGG----------SDMLPAQQ----------KMVPLPF 769

Qy       1712 GLSPRESSLALNYAAGPRGIIDLSQVP----HLPVLVPPTPGTPATAMDRLAYLPTAPQP 1767
Db        770 GEHPQQ-----EYGMGPRPFLPMSQGPGSNSGLRNLREPI-GPDQRTNSRLSHMPPLPLN 823

Qy       1768 FSSRHSSSPLSPGGPTHLTKPTTTSSSERERDRDRERDRDREREKSILTSTTTVEHAPIW 1827
Db        824 PSSNPTSLNTAP--------PVQRGLGRKPLD------------------ISVAGSQVH 856

Qy       1828 RPGTEQSSGSSGSSGGGGGSSSRPASHSHAHQ-HSPI--SPRTQDALQQRPSVLHNTGM- 1883
Db        857 SPGI-----------------NPLKSPTMHQVQSPMLGSPSGNLKSPQTPSQL--AGML 896

Qy       1884 KGIITAVEPSKPTVLRSTSTSSPVRPAATFPPATHCPLGGTLDGVYPTLMEPVLLPK-EA 1942
Db        897 AGPAAAASIKSPPVLGSAAAS----PVHLKSPSLPAPSPGWTSSPKPPLQSPGIPPNHKA 952

Qy       1943 P-RVARPERPRADTGHAFLAKPPARSGLEPASSPSKGSEPRPLVPPVSGHATIARTP--- 1998
Db        953 PLTMASP----AMLGNVESGGPPPPTASQPASVNIPGSLPSSTPYTMPPEPTLSQNPLSI 1008

Qy       1999 -----AKNLAP-----HH------ASPDPPAPPASASDPHREKTQSKPFSIQELELRSLG 2042
Db       1009 MMSRMSKFAMPSSTPLYHDAIKTVASSDDDSPPARS--PNLPSMNNMP---------GMG 1057
```



RESULT 93
US-10-322-579-15
; Sequence 15, Application US/10322579
; Publication No. US20030114413A1
; GENERAL INFORMATION:
; APPLICANT: BASLER, Konrad
; APPLICANT: BRUNNER, Erich
; APPLICANT: FROESCH, Barbara
; APPLICANT: KRAMPS, Thomas
; APPLICANT: PETER, Oliver
; TITLE OF INVENTION: ESSENTIAL DOWNSTREAM COMPONENT OF THE WINGLESS SIGNALING PATHWAY
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC APPLICATIONS BASED THEREON
; FILE REFERENCE: Q60361
; CURRENT APPLICATION NUMBER: US/10/322,579
; CURRENT FILING DATE: 2002-12-19
; PRIOR APPLICATION NUMBER: US/09/915,543
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 60/221,502
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 1426
; TYPE: PRT
; ORGANISM: Human lgs/bcl9
US-10-322-579-15

Query Match 2.8%; Score 364; DB 14; Length 1426;
Best Local Similarity 19.7%; Pred. No. 2.2e-08;
Matches 377; Conservative 197; Mismatches 685; Indels 658; Gaps 98;



```
Db        857 SPGI------------------NPLKSPTMHQVQSPMLGSPSGNLKSPQTPSQL--AGML 896

Qy       1884 KGIITAVEPSKPTVLRSTSTSSPVRPAATFPPATHCPLGGTLDGVYPTLMEPVLLPK-EA 1942

Db        897 AGPAAAASIKSPPVLGSAAAS----PVHLKSPSLPAPSPGWTSSPKPPLQSPGIPPNHKA 952

Qy       1943 P-RVARPERPRADTGHAFLAKPPARSGLEPASSPSKGSEPRPLVPPVSGHATIARTP--- 1998

Db        953 PLTMASP----AMLGNVESGGPPPPTASQPASVNIPGSLPSSTPYTMPPEPTLSQNPLSI 1008

Qy       1999 -----AKNLAP-----HH------ASPDPPAPPASASDPHREKTQSKPFSIQELELRSLG 2042

Db       1009 MMSRMSKFAMPSSTPLYHDAIKTVASSDDDSPPARS--PNLPSMNNMP---------GMG 1057

Qy       2043 YHGSSYSPEGVEPVSPVSSPSLTHDKGLPKHLEELDKSHLEGELRPKQPGPVKLGGEAAH 2102

Db       1058 INTQNPRISGPNPVVPM--PTLS-PMGMTQPL-----SHSN-----QMPSPNAVG----- 1099

Qy       2103 LPHLRPLPESQPSSSPLLQTAPGVKGHQRVVTLAQHISEVITQDYTRHHPQQLSAPLPAP 2162

Db       1100 -------PNIPPHGVPM---GPGLMSHNPIM---------------GHGSQEPPMVPQG 1133

Qy       2163 LYSFPGASCPVLDLRRPPSDLYLPPPDHGAPARGSPHSEGG------------KRSPEPN 2210

Db       1134 RMGFPQGFPPV---QSPPQQVPFP---HNGPSGGQGSFPGGMGFPGEGPLGRPSNLPQSS 1187

Qy       2211 KTSVL--GGGEDGIEPVSPPEGMTEPGHSRSAVYPLLYRDGEQTEPSRMGSKSPGNTSQP 2268

Db       1188 ADAALCKPGGPGG------PDSFTVLGNSMPSVF---------TDPDLQEVIRPGATGIP 1232

Qy       2269 PAFFSKLTESNSAMVKSKKQEINKKLNTHNRNEPEYNISQPGTEIFNMPAITGTGLMTYR 2328

Db       1233 EFDLSRIIPSEK---PSQTLQYFPRGEVPGRKQPQ----GPGPGFSHMQGMMG------- 1278

Qy       2329 SQAVQEHASTNMGLEAIIRKALMGKYDQWEESPPLSANAFNPLNASASLPAAMPIT---- 2384

Db       1279 ------EQAPRMGL------ALPGM-----GGPGPVGTPDIPLGTAPSMPGHNPMRPPAF 1321

Qy       2385 AADGR--SDHTLTSPGGGGKAKVSGRPSSRKAKSPAPGLASG-DRPPSVSSVHSEGDCNR 2441

Db       1322 LQQGMMGPHHRMMSP---AQSTMPGQPTLMSNPAAAVGMIPGKDRGPAGLYTH------- 1371

Qy       2442 RTPLTNRVWEDRPSSAGSTPFPYNPLIMRLQAGVMAS------PP---PPGLPAGSG 2489

Db       1372 ------------PGPVGS------PGMMMSMQGMMGPQQNIMIPPQMRPRGMAADVG 1410
```

```
RESULT 94
US-09-963-875-1
; Sequence 1, Application US/09963875
; Patent No. US20020164307A1
; GENERAL INFORMATION:
;  APPLICANT: Massachusetts General Hospital
;  TITLE OF INVENTION: Stem Cells of the Islets of Langerhans and Their Use in Treating
;  TITLE OF INVENTION:  Mellitus
;  FILE REFERENCE: 17633/1235
;  CURRENT APPLICATION NUMBER: US/09/963,875
;  CURRENT FILING DATE:  2001-09-26
;  PRIOR APPLICATION NUMBER: US60/169082
;  PRIOR FILING DATE: 1999-12-06
;  PRIOR APPLICATION NUMBER: US 60/215109
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: US 60/238880
;  PRIOR FILING DATE: 2000-10-06
;  PRIOR APPLICATION NUMBER: US 09/731261
;  PRIOR FILING DATE: 2000-12-06
;  NUMBER OF SEQ ID NOS: 58
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 1618
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-963-875-1
```

```
  Query Match             2.7%;  Score 358;  DB 9;  Length 1618;
  Best Local Similarity   21.1%;  Pred. No. 4.9e-08;
  Matches  323;  Conservative  196;  Mismatches  511;  Indels  502;  Gaps   77;

Qy         97 KSEMEFIESKRPRL--ELLP--DPLLRPSPL---LATGQPA--GSEDLTKDRSLTGKLEP 147

Db        329 KLELQFPRTPEGRRLGSLLPVLSPTSLPSPLPATLETPVPAFLKNQEFLQARTPTLASTP 388

Qy        148 VSPPSPPHTDPELELVPPRLSKEELIQNMDRVDREITMVEQQISKL------KKKQQQLE 201

Db        389 I-PPTPQAPSP----------------AVDAEIRAQDAPLSLLQTQGGRKQAPEPLR 428

Qy        202 EEA--AKP----PEPEKP----------VSP-------PPIESKH--------------- 223

Db        429 AEARVAIPASVLPGPEEPGGQRQEASTGQSPEDHASLAPPLSPDHSSLEAKDGESGGSRV 488

Qy        224 --------------------------RSLVQIIYDE---NRKKAEAAH-----RILEGL 248

Db        489 FSICRGEGEGQIWGLVEKETAIEGKVVSSLQQEIWEEEDLNRKEIQDSQVPLEKETLKSL 548

Qy        249 GPQVE---LPLYNQPSDTRQYHENIKINQAMRKKLILYFKRRNHARKQWKQKFCQRYDQL 305

Db        549 GEEIQESLKTLENQSHETLE-RENQECPRSLEEDL----------------------ET 584

Qy        306 MEALEKKVERI-------ENNPRRRAKESKVREYYEKQFPEIRKQRELQERMQSRVGQRG 358

Db        585 LKSLEKENKRAIKGCGGSETSRKRGCRQLKPTGKEDTQTLQ-SLQKENQELMKSLEGNLE 643

Qy        359 SGLSMSAARSEHEVSEIIDGLSEQENLEKQMRQLAVIPPMLYDADQQRIKFINMNGLMAD 418

Db        644 TFL-FPGTENQELVSSLQENLESLTALEKENQEPLRSPEV---GDEEALRPLTKEN--QE 697

Qy        419 PMKVYKDRQVMNMWSEQEKETFR--EKFMQHPKNFGLIASFLERKTVAECVLYYYLTKKN 476

Db        698 PLRSLED---------ENKEAFRSLEKENQEP----------LKTLEEEDQSIVRPLET 737

Qy        477 ENYKSL------VRRSYRRRGKSQQQQQQQQQQQQQQQQQPMPRSSQEEKDEKEKEKEAE 530

Db        738 ENHKSLRSLEEQDQETLRTLEKETQQRRRSLGEQDQMTLRP------PEKVDLEPLKSLD 791

Qy        531 KEEEKPEVENDKEDLLKEKTDDTSGEDNDEKEAVASKGRKTANSQGRRKGRITRSMANEA 590

Db        792 QEIARP-LENENQEFLKSLKEES-------VEAVKSLETEILES---------LKSAGQE 834

Qy        591 NSEEAITPQQSAELASMELNESSRWTEEEM------ETAKKG-------------LLEH 630

Db        835 NLETLKSPETQAPL----------WTPEEINKSGGNESSRKGNSRTTGVCGSEPRDIQTP 884

Qy        631 GRNWSAIARMVGSKTVSQC-------KNFYFNYKKRQNLDEILQQHKLKMEKERNARRKK 683

Db        885 GRGESGIIEISGSMEPGEFEISRGVDKESQRNLEEEENLGKGEYQESL-----RSLEEEG 939

Qy        684 KKAPAAASEEAAFPPVVEDEEME------ASGVSGNEEEMVEEAEALHASGNE--VPRGE 735

Db        940 QELPQSADVQRWEDTVEKDQELAQESPPGMAGVENKDEAELNLREQDGFTGKEEVVEQGE 999

Qy        736 CSGPATVNNSSDTESIPSP---HTEAAKDTGQNGPKPPATL--GADGPPPGPPTPPRRTS 790

Db       1000 LNA---------TEEVWFPGEGHPENPEPKEQRGLVEGASVKGGAEGLQ---------- 1039

Qy        791 RAPIEPTPASEATGAPTPPPAPPSPSAPPPVVPKEEKEEETAAAPPVEEGEEQKPPAAEE 850

Db       1040 ----DPEGQSQQVGTPGLQAPQGLPEAIEPLVEDDVAPGGDQASPEVMLGSE---PAMGE 1092

Qy        851 LAV-------------DTGK------AEEPVKSECTE----EAEEGPAKGKDAEAAEAT 886

Db       1093 SAAGAEPGLGQGVGGLGDPGHLTREEVMEPPLEEESLEAKRVQGLEGPR--KDLEEA--- 1147

Qy        887 AEGALKAEKKE-GGSGRATTAKSSGAPQDSDSSATCSADEVDEAE-----GGDKNRLLSP 940

Db       1148 --GGLGTEFSELPGKSRDPWEPPREGREESEAEAPRGAEEAFPAETLGHTGSD-----AP 1200
```

```
Qy        941 RPSLLTPTGDPRANASPQKPLDLKQLKQRAAAIPPIQV--TKVHEPPREDAAPTKPAPPA 998
Db       1201 SP---WPLGSEEAEED---------------VPPVLVSPSPTYTPILEDA--------- 1232

Qy        999 PPPPQNLQPESDAPQQPGSSPRGKSRSPAPPADKEAFAAEAQKLPGDPPCWTSGLPFPVP 1058
Db       1233 ----PGLQPQAEGSQEASWGVQGRAEAGKVESEQEELGS------GEIP---EGL----- 1274

Qy       1059 PREVIKASPHAPDPSAFSYAPPGHPLPLGLHDTARPVLPR-PPTISNPPPLISSAKH--- 1114
Db       1275 QEEGEESREESEEDELGETLPDSTPLGFYLRS---PTSPRWTPLESRGHPLKETGKEGWD 1331

Qy       1115 PSVL------------ERQIGAISQGMSVQLHVPYSE-HAKAPVGPVTMGLPLPMDPKKL 1161
Db       1332 PAVLASEGLEEPSEKEEGEEGEEECGRDSDLSEEFEDLGTEAPF--------LPGVPGEV 1383

Qy       1162 A-PFSGVKQEQLSP--------------------------------RGQAGPPESLG-- 1185
Db       1384 AEPLGQVPQLLLDPAAWDRDGESDGFADEEESGEEGEEDQEEGREPGAGRWGPGSSVGSL 1443

Qy       1186 --VPTAQEASVLRGTALG-SVP-----GGSITKGIPSTRVPSDS--AITYRGSITHGTPA 1235
Db       1444 QALSSSQRGEFLESDSVSVSVPWDDSLRGAVA-GAPKTALETESQDSAEPSGSEEESDPV 1502

Qy       1236 DVLYKGTITRIIGEDSPSRLDRGREDSLPKGHVI--------YEGKKGHVLSYEGG-MSV 1286
Db       1503 SLEREDKV------PGPLEIPSGMEDAGPGADIIGVNGQGPNLEGKSQHV---NGGVMNG 1553

Qy       1287 TQCSKEDG----------RSSSGPPHETAAPKRTYDMMEGRVGRAISSASIEGLMGRAIP 1336
Db       1554 LEQSEESGARNALVSEGDRGSPFQEEEGSALKR-------------SSAG---------- 1590

Qy       1337 PERHSPHHLKEQHHIRGSITQGIPRSYVEAQE 1368
Db       1591 ----APVHLGQGQFLKFTQREGDRESWSSGED 1618


RESULT 95
US-09-731-255-2
; Sequence 2, Application US/09731255
; Publication No. US20010024824A1
; GENERAL INFORMATION:
;  APPLICANT: Habener, Joel
;  APPLICANT:  Zulewski, Hendrik
;  APPLICANT:  Abraham, Elizabeth
;  APPLICANT:  Vallejo, Mario
;  TITLE OF INVENTION: STEM CELLS OF THE ISLETS OF LANGERHANS AND THEIR USE IN TREATING
;  TITLE OF INVENTION:  DIABETES MELLITUS
;  FILE REFERENCE: 3284/1220
;  CURRENT APPLICATION NUMBER: US/09/731,255
;  CURRENT FILING DATE:  2000-12-06
;  PRIOR APPLICATION NUMBER: US 60/169,082
;  PRIOR FILING DATE: 1999-12-06
;  PRIOR APPLICATION NUMBER: US 60/215,109
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: US 60/239,880
;  PRIOR FILING DATE: 2000-10-06
;  NUMBER OF SEQ ID NOS: 55
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
;   LENGTH: 1618
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-731-255-2

  Query Match             2.7%;  Score 358;  DB 12;  Length 1618;
  Best Local Similarity   21.1%;  Pred. No. 4.9e-08;
  Matches  323;  Conservative  196;  Mismatches  511;  Indels  502;  Gaps   77;

Qy         97 KSEMEFIESKRPRL--ELLP--DPLLRPSPL---LATGQPA--GSEDLTKDRSLTGKLEP 147
Db        329 KLELQFPRTPEGRRLGSLLPVLSPTSLPSPLPATLETPVPAFLKNQEFLQARTPTLASTP 388

Qy        148 VSPPSPPHTDPELELVPPRLSKEELIQNMDRVDREITMVEQQISKL------KKKQQQLE 201
Db        389 I-PPTPQAPSP------------------AVDAEIRAQDAPLSLLQTQGGRKQAPEPLR 428

Qy        202 EEA--AKP----PEPEKP----------VSP-------PPIESKH--------------- 223
Db        429 AEARVAIPASVLPGPEEPGGQRQEASTGQSPEDHASLAPPLSPDHSSLEAKDGESGGSRV 488

Qy        224 --------------------------RSLVQIIYDE---NRKKAEAAH-----RILEGL 248
Db        489 FSICRGEGEGQIWGLVEKETAIEGKVVSSLQQEIWEEEDLNRKEIQDSQVPLEKETLKSL 548

Qy        249 GPQVE---LPLYNQPSDTRQYHENIKINQAMRKKLILYFKRRNHARKQWKQKFCQRYDQL 305
Db        549 GEEIQESLKTLENQSHETLE-RENQECPRSLEEDL----------------------ET 584

Qy        306 MEALEKKVERI-------ENNPRRRAKESKVREYYEKQFPEIRKQRELQERMQSRVGQRG 358
Db        585 LKSLEKENKRAIKGCGGSETSRKRGCRQLKPTGKEDTQTLQ-SLQKENQELMKSLEGNLE 643

Qy        359 SGLSMSAARSEHEVSEIIDGLSEQENLEKQMRQLAVIPPMLYDADQQRIKFINMNGLMAD 418
Db        644 TFL-FPGTENQELVSSLQENLESLTALEKENQEPLRSPEV---GDEEALRPLTKEN--QE 697

Qy        419 PMKVYKDRQVMNMWSEQEKETFR--EKFMQHPKNFGLIASFLERKTVAECVLYYYLTKKN 476
Db        698 PLRSLED---------ENKEAFRSLEKENQEP----------LKTLEEEDQSIVRPLET 737

Qy        477 ENYKSL------VRRSYRRRGKSQQQQQQQQQQQQQQQQQQPMPRSSQEEKDEKEKEKEAE 530
Db        738 ENHKSLRSLEEQDQETLRTLEKETQQRRRSLGEQDQMTLRP------PEKVDLEPLKSLD 791

Qy        531 KEEEKPEVENDKEDLLKEKTDDTSGEDNDEKEAVASKGRKTANSQGRRKGRITRSMANEA 590
Db        792 QEIARP-LENENQEFLKSLKEES-------VEAVKSLETEILES---------LKSAGQE 834

Qy        591 NSEEAITPQQSAELASMELNESSRWTEEEM------ETAKKG--------------LLEH 630
Db        835 NLETLKSPETQAPL----------WTPEEINKSGGNESSRKGNSRTTGVCGSEPRDIQTP 884

Qy        631 GRNWSAIARMVGSKTVSQC-------KNFYFNYKKRQNLDEILQQHKLKMEKERNARRKK 683
Db        885 GRGESGIIEISGSMEPGEFEISRGVDKESQRNLEEEENLGKGEYQESL-----RSLEEEG 939

Qy        684 KKAPAAASEEAAFPPVVEDEEME------ASGVSGNEEEMVEEAEALHASGNE--VPRGE 735
Db        940 QELPQSADVQRWEDTVEKDQELAQESPPGMAGVENKDEAELNLREQDGFTGKEEVVEQGE 999

Qy        736 CSGPATVNNSSDTESIPSP---HTEAAKDTGQNGPKPPATL--GADGPPPGPPTPPRRTS 790
Db       1000 LNA---------TEEVWFPGEGHPENPEPKEQRGLVEGASVKGGAEGLQ----------- 1039

Qy        791 RAPIEPTPASEATGAPTPPPAPPSPSAPPPVVPKEEKEEETAAAPPVEEGEEQKPPAAEE 850
Db       1040 ----DPEGQSQQVGTPGLQAPQGLPEAIEPLVEDDVAPGGDQASPEVMLGSE---PAMGE 1092

Qy        851 LAV-------------DTGK------AEEPVKSECTE----EAEEGPAKGKDAEAAEAT 886
Db       1093 SAAGAEPGLGQGVGGLGDPGHLTREEVMEPPLEEESLEAKRVQGLEGPR--KDLEEA--- 1147

Qy        887 AEGALKAEKKE-GGSGRATTAKSSGAPQDSDSSATCSADEVDEAE-----GGDKNRLLSP 940
Db       1148 --GGLGTEFSELPGKSRDPWEPPREGREESEAEAPRGAEEAFPAETLGHTGSD-----AP 1200

Qy        941 RPSLLTPTGDPRANASPQKPLDLKQLKQRAAAIPPIQV--TKVHEPPREDAAPTKPAPPA 998
Db       1201 SP---WPLGSEEAEED---------------VPPVLVSPSPTYTPILEDA--------- 1232

Qy        999 PPPPQNLQPESDAPQQPGSSPRGKSRSPAPPADKEAFAAEAQKLPGDPPCWTSGLPFPVP 1058
Db       1233 ----PGLQPQAEGSQEASWGVQGRAEAGKVESEQEELGS------GEIP---EGL----- 1274

Qy       1059 PREVIKASPHAPDPSAFSYAPPGHPLPLGLHDTARPVLPR-PPTISNPPPLISSAKH--- 1114
```

```
Db       1275 QEEGEESREESEEDELGETLPDSTPLGFYLRS---PTSPRWTPLESRGHPLKETGKEGWD 1331

Qy       1115 PSVL------------ERQIGAISQGMSVQLHVPYSE-HAKAPVGPVTMGLPLPMDPKKL 1161

Db       1332 PAVLASEGLEEPSEKEEGEEGEEECGRDSDLSEEFEDLGTEAPF--------LPGVPGEV 1383

Qy       1162 A-PFSGVKQEQLSP--------------------------------RGQAGPPESLG-- 1185

Db       1384 AEPLGQVPQLLLDPAAWDRDGESDGFADEEESGEEGEEDQEEGREPGAGRWGPGSSVGSL 1443

Qy       1186 --VPTAQEASVLRGTALG-SVP-----GGSITKGIPSTRVPSDS--AITYRGSITHGTPA 1235

Db       1444 QALSSSQRGEFLESDSVSVSVPWDDSLRGAVA-GAPKTALETESQDSAEPSGSEEESDPV 1502

Qy       1236 DVLYKGTITRIIGEDSPSRLDRGREDSLPKGHVI--------YEGKKGHVLSYEGG-MSV 1286

Db       1503 SLEREDKV------PGPLEIPSGMEDAGPGADIIGVNGQGPNLEGKSQHV---NGGVMNG 1553

Qy       1287 TQCSKEDG----------RSSSGPPHETAAPKRTYDMMEGRVGRAISSASIEGLMGRAIP 1336

Db       1554 LEQSEESGARNALVSEGDRGSPFQEEEGSALKR------------SSAG---------- 1590

Qy       1337 PERHSPHHLKEQHHIRGSITQGIPRSYVEAQE 1368

Db       1591 ----APVHLGQGQFLKFTQREGDRESWSSGED 1618


RESULT 96
US-10-136-891-2
; Sequence 2, Application US/10136891
; Publication No. US20030031657A1
; GENERAL INFORMATION:
;  APPLICANT: Habener, Joel
;  TITLE OF INVENTION: STEM CELLS AND THEIR USE IN TRANSPLANTATION
;  FILE REFERENCE: 3284/1225
;  CURRENT APPLICATION NUMBER: US/10/136,891
;  CURRENT FILING DATE:  2002-10-01
;  PRIOR APPLICATION NUMBER: US 60/169,082
;  PRIOR FILING DATE: 1999-12-06
;  PRIOR APPLICATION NUMBER: US 60/215,109
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: US 60/239,880
;  PRIOR FILING DATE: 2000-10-06
;  PRIOR APPLICATION NUMBER: US 09/731,255
;  PRIOR FILING DATE: 2000-12-06
;  NUMBER OF SEQ ID NOS: 55
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 1618
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-136-891-2

  Query Match             2.7%;  Score 358;  DB 14;  Length 1618;
  Best Local Similarity   21.1%;  Pred. No. 4.9e-08;
  Matches  323;  Conservative  196;  Mismatches  511;  Indels  502;  Gaps   77;

Qy          97 KSEMEFIESKRPRL--ELLP--DPLLRPSPL---LATGQPA--GSEDLTKDRSLTGKLEP 147

Db         329 KLELQFPRTPEGRRLGSLLPVLSPTSLPSPLPATLETPVPAFLKNQEFLQARTPTLASTP 388

Qy         148 VSPPSPPHTDPELELVPPRLSKEELIQNMDRVDREITMVEQQISKL------KKKQQQLE 201

Db         389 I-PPTPQAPSP------------------AVDAEIRAQDAPLSLLQTQGGRKQAPEPLR 428

Qy         202 EEA--AKP----PEPEKP----------VSP-------PPIESKH--------------- 223

Db         429 AEARVAIPASVLPGPEEPGGQRQEASTGQSPEDHASLAPPLSPDHSSLEAKDGESGGSRV 488

Qy         224 --------------------------RSLVQIIYDE---NRKKAEAAH-----RILEGL 248

Db         489 FSICRGEGEGQIWGLVEKETAIEGKVVSSLQQEIWEEEDLNRKEIQDSQVPLEKETLKSL 548


Qy         249 GPQVE---LPLYNQPSDTRQYHENIKINQAMRKKLILYFKRRNHARKQWKQKFCQRYDQL 305

Db         549 GEEIQESLKTLENQSHETLE-RENQECPRSLEEDL----------------------ET 584

Qy         306 MEALEKKVERI-------ENNPRRRAKESKVREYYEKQFPEIRKQRELQERMQSRVGQRG 358

Db         585 LKSLEKENKRAIKGCGGSETSRKRGCRQLKPTGKEDTQTLQ-SLQKENQELMKSLEGNLE 643

Qy         359 SGLSMSAARSEHEVSEIIDGLSEQENLEKQMRQLAVIPPMLYDADQQRIKFINMNGLMAD 418

Db         644 TFL-FPGTENQELVSSLQENLESLTALEKENQEPLRSPEV---GDEEALRPLTKEN--QE 697

Qy         419 PMKVYKDRQVMNMWSEQEKETFR--EKFMQHPKNFGLIASFLERKTVAECVLYYYLTKKN 476

Db         698 PLRSLED---------ENKEAFRSLEKENQEP----------LKTLEEEDQSIVRPLET 737

Qy         477 ENYKSL------VRRSYRRRGKSQQQQQQQQQQQQQQQQQPMPRSSQEEKDEKEKEKEAE 530

Db         738 ENHKSLRSLEEQDQETLRTLEKETQQRRRSLGEQDQMTLRP------PEKVDLEPLKSLD 791

Qy         531 KEEEKPEVENDKEDLLKEKTDDTSGEDNDEKEAVASKGRKTANSQGRRKGRITRSMANEA 590

Db         792 QEIARP-LENENQEFLKSLKEES-------VEAVKSLETEILES---------LKSAGQE 834

Qy         591 NSEEAITPQQSAELASMELNESSRWTEEEM------ETAKKG--------------LLEH 630

Db         835 NLETLKSPETQAPL----------WTPEEINKSGGNESSRKGNSRTTGVCGSEPRDIQTP 884

Qy         631 GRNWSAIARMVGSKTVSQC-------KNFYFNYKKRQNLDEILQQHKLKMEKERNARRKK 683

Db         885 GRGESGIIEISGSMEPGEFEISRGVDKESQRNLEEEENLGKGEYQESL-----RSLEEEG 939

Qy         684 KKAPAAAASEEAAFPPVVEDEEME------ASGVSGNEEEMVEEAEALHASGNE--VPRGE 735

Db         940 QELPQSADVQRWEDTVEKDQELAQESPPGMAGVENKDEAELNLREQDGFTGKEEVVEQGE 999

Qy         736 CSGPATVNNSSDTESIPSP---HTEAAKDTGQNGPKPPATL--GADGPPPGPPTPPRRTS 790

Db        1000 LNA---------TEEVWFPGEGHPENPEPKEQRGLVEGASVKGGAEGLQ---------- 1039

Qy         791 RAPIEPTPASEATGAPTPPPAPPSPSAPPPVVPKEEKEEETAAAPPVEEGEEQKPPAAEE 850

Db        1040 ----DPEGQSQQVGTPGLQAPQGLPEAIEPLVEDDVAPGGDQASPEVMLGSE---PAMGE 1092

Qy         851 LAV--------------DTGK------AEEPVKSECTE----EAEEGPAKGKDAEAAEAT 886

Db        1093 SAAGAEPGLGQGVGGLGDPGHLTREEVMEPPLEEESLEAKRVQGLEGPR--KDLEEA--- 1147

Qy         887 AEGALKAEKKE-GGSGRATTAKSSGAPQDSDSSATCSADEVDEAE-----GGDKNRLLSP 940

Db        1148 --GGLGTEFSELPGKSRDPWEPPREGREESEAEAPRGAEEAFPAETLGHTGSD-----AP 1200

Qy         941 RPSLLTPTGDPRANASPQKPLDLKQLKQRAAAIPPIQV--TKVHEPPREDAAPTKPAPPA 998

Db        1201 SP---WPLGSEEAEED---------------VPPVLVSPSPTYTPILEDA--------- 1232

Qy         999 PPPPQNLQPESDAPQQPGSSPRGKSRSPAPPADKEAFAAEAQKLPGDPPCWTSGLPFPVP 1058

Db        1233 ----PGLQPQAEGSQEASWGVQGRAEAGKVESEQEELGS------GEIP---EGL----- 1274

Qy        1059 PREVIKASPHAPDPSAFSYAPPGHPLPLGLHDTARPVLPR-PPTISNPPPLISSAKH--- 1114

Db        1275 QEEGEESREESEEDELGETLPDSTPLGFYLRS---PTSPRWTPLESRGHPLKETGKEGWD 1331

Qy        1115 PSVL------------ERQIGAISQGMSVQLHVPYSE-HAKAPVGPVTMGLPLPMDPKKL 1161

Db        1332 PAVLASEGLEEPSEKEEGEEGEEECGRDSDLSEEFEDLGTEAPF--------LPGVPGEV 1383

Qy        1162 A-PFSGVKQEQLSP--------------------------------RGQAGPPESLG-- 1185

Db        1384 AEPLGQVPQLLLDPAAWDRDGESDGFADEEESGEEGEEDQEEGREPGAGRWGPGSSVGSL 1443
```

```
Qy       1186 --VPTAQEASVLRGTALG-SVP-----GGSITKGIPSTRVPSDS--AITYRGSITHGTPA 1235
Db       1444 QALSSSQRGEFLESDSVSVSVPWDDSLRGAVA-GAPKTALETESQDSAEPSGSEEESDPV 1502

Qy       1236 DVLYKGTITRIIGEDSPSRLDRGREDSLPKGHVI--------YEGKKGHVLSYEGG-MSV 1286
Db       1503 SLEREDKV------PGPLEIPSGMEDAGPGADIIGVNGQGPNLEGKSQHV---NGGVMNG 1553

Qy       1287 TQCSKEDG----------RSSSGPPHETAAPKRTYDMMEGRVGRAISSASIEGLMGRAIP 1336
Db       1554 LEQSEESGARNALVSEGDRGSPFQEEEGSALKR------------SSAG---------- 1590

Qy       1337 PERHSPHHLKEQHHIRGSITQGIPRSYVEAQE 1368
Db       1591 ----APVHLGQGQFLKFTQREGDRESWSSGED 1618
```

RESULT 97
US-10-120-687-1
; Sequence 1, Application US/10120687
; Publication No. US20030082155A1
; GENERAL INFORMATION:
; APPLICANT: Massachusetts General Hospital
; TITLE OF INVENTION: Stem Cells of the Islets of Langerhans and Their Use in Treating
; TITLE OF INVENTION: Mellitus
; FILE REFERENCE: 3284/1235B
; CURRENT APPLICATION NUMBER: US/10/120,687
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: US60/169082
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: US 09/963,875
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/215109
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: US 60/238880
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 09/731261
; PRIOR FILING DATE: 2000-12-06
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1618
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-120-687-1

Query Match 2.7%; Score 358; DB 14; Length 1618;
Best Local Similarity 21.1%; Pred. No. 4.9e-08;
Matches 323; Conservative 196; Mismatches 511; Indels 502; Gaps 77;

```
Qy         97 KSEMEFIESKRPRL--ELLP--DPLLRPSPL---LATGQPA--GSEDLTKDRSLTGKLEP 147
Db        329 KLELQFPRTPEGRRLGSLLPVLSPTSLPSPLPATLETPVPAFLKNQEFLQARTPTLASTP 388

Qy        148 VSPPSPPHTDPELELVPPRLSKEELIQNMDRVDREITMVEQQISKL------KKKQQQLE 201
Db        389 I-PPTPQAPSP------------------AVDAEIRAQDAPLSLLQTQGGRKQAPEPLR 428

Qy        202 EEA--AKP----PEPEKP----------VSP-------PPIESKH--------------- 223
Db        429 AEARVAIPASVLPGPEEPGGQRQEASTGQSPEDHASLAPPLSPDHSSLEAKDGESGGSRV 488

Qy        224 --------------------------RSLVQIIYDE---NRKKAEAAH-----RILEGL 248
Db        489 FSICRGEGEGQIWGLVEKETAIEGKVVSSLQQEIWEEEDLNRKEIQDSQVPLEKETLKSL 548

Qy        249 GPQVE---LPLYNQPSDTRQYHENIKINQAMRKKLILYFKRRNHARKQWKQKFCQRYDQL 305
Db        549 GEEIQESLKTLENQSHETLE-RENQECPRSLEEDL----------------------ET 584

Qy        306 MEALEKKVERI-------ENNPRRRAKESKVREYYEKQFPEIRKQRELQERMQSRVGQRG 358
Db        585 LKSLEKENKRAIKGCGGSETSRKRGCRQLKPTGKEDTQTLQ-SLQKENQELMKSLEGNLE 643

Qy        359 SGLSMSAARSEHEVSEIIDGLSEQENLEKQMRQLAVIPPMLYDADQQRIKFINMNGLMAD 418
Db        644 TFL-FPGTENQELVSSLQENLESLTALEKENQEPLRSPEV---GDEEALRPLTKEN--QE 697

Qy        419 PMKVYKDRQVMNMWSEQEKETFR--EKFMQHPKNFGLIASFLERKTVAECVLYYYLTKKN 476
Db        698 PLRSLED---------ENKEAFRSLEKENQEP----------LKTLEEEDQSIVRPLET 737

Qy        477 ENYKSL------VRRSYRRRGKSQQQQQQQQQQQQQQQQQPMPRSSQEEKDEKEKEKEAE 530
Db        738 ENHKSLRSLEEQDQETLRTLEKETQQRRRSLGEQDQMTLRP------PEKVDLEPLKSLD 791

Qy        531 KEEEKPEVENDKEDLLKEKTDDTSGEDNDEKEAVASKGRKTANSQGRRKGRITRSMANEA 590
Db        792 QEIARP-LENENQEFLKSLKEES-------VEAVKSLETEILES---------LKSAGQE 834

Qy        591 NSEEAITPQQSAELASMELNESSRWTEEEM------ETAKKG--------------LLEH 630
Db        835 NLETLKSPETQAPL----------WTPEEINKSGGNESSRKGNSRTTGVCGSEPRDIQTP 884

Qy        631 GRNWSAIARMVGSKTVSQC-------KNFYFNYKKRQNLDEILQQHKLKMEKERNARRKK 683
Db        885 GRGESGIIEISGSMEPGEFEISRGVDKESQRNLEEEENLGKGEYQESL-----RSLEEEG 939

Qy        684 KKAPAAASEEAAFPPVVEDEEME------ASGVSGNEEEMVEEAEALHASGNE--VPRGE 735
Db        940 QELPQSADVQRWEDTVEKDQELAQESPPGMAGVENKDEAELNLREQDGFTGKEEVVEQGE 999

Qy        736 CSGPATVNNSSDTESIPSP---HTEAAKDTGQNGPKPPATL--GADGPPPGPPTPPRRTS 790
Db       1000 LNA---------TEEVWFPGEGHPENPEPKEQRGLVEGASVKGGAEGLQ---------- 1039

Qy        791 RAPIEPTPASEATGAPTPPPAPPSPSAPPPVVPKEEKEEETAAAPPVEEGEEQKPPAAEE 850
Db       1040 ----DPEGQSQQVGTPGLQAPQGLPEAIEPLVEDDVAPGGDQASPEVMLGSE---PAMGE 1092

Qy        851 LAV--------------DTGK------AEEPVKSECTE----EAEEGPAKGKDAEAAEAT 886
Db       1093 SAAGAEPGLGQGVGGLGDPGHLTREEVMEPPLEEESLEAKRVQGLEGPR--KDLEEA--- 1147

Qy        887 AEGALKAEKKE-GGSGRATTAKSSGAPQDSDSSATCSADEVDEAE-----GGDKNRLLSP 940
Db       1148 --GGLGTEFSELPGKSRDPWEPPREGREESEAEAPRGAEEAFPAETLGHTGSD-----AP 1200

Qy        941 RPSLLTPTGDPRANASPQKPLDLKQLKQRAAAIPPIQV--TKVHEPPREDAAPTKPAPPA 998
Db       1201 SP---WPLGSEEAEED---------------VPPVLVSPSPTYTPILEDA--------- 1232

Qy        999 PPPPQNLQPESDAPQQPGSSPRGKSRSPAPPADKEAFAAEAQKLPGDPPCWTSGLPFPVP 1058
Db       1233 ----PGLQPQAEGSQEASWGVQGRAEAGKVESEQEELGS------GEIP---EGL----- 1274

Qy       1059 PREVIKASPHAPDPSAFSYAPPGHPLPLGLHDTARPVLPR-PPTISNPPPLISSAKH--- 1114
Db       1275 QEEGEESREESEEDELGETLPDSTPLGFYLRS---PTSPRWTPLESRGHPLKETGKEGWD 1331

Qy       1115 PSVL------------ERQIGAISQGMSVQLHVPYSE-HAKAPVGPVTMGLPLPMDPKKL 1161
Db       1332 PAVLASEGLEEPSEKEEGEEGEEECGRDSDLSEEFEDLGTEAPF--------LPGVPGEV 1383

Qy       1162 A-PFSGVKQEQLSP--------------------------------RGQAGPPESLG-- 1185
Db       1384 AEPLGQVPQLLLDPAAWDRDGESDGFADEEESGEEGEEDQEEGREPGAGRWGPGSSVGSL 1443

Qy       1186 --VPTAQEASVLRGTALG-SVP-----GGSITKGIPSTRVPSDS--AITYRGSITHGTPA 1235
Db       1444 QALSSSQRGEFLESDSVSVSVPWDDSLRGAVA-GAPKTALETESQDSAEPSGSEEESDPV 1502

Qy       1236 DVLYKGTITRIIGEDSPSRLDRGREDSLPKGHVI--------YEGKKGHVLSYEGG-MSV 1286
Db       1503 SLEREDKV------PGPLEIPSGMEDAGPGADIIGVNGQGPNLEGKSQHV---NGGVMNG 1553
```

```
Qy       1287 TQCSKEDG----------RSSSGPPHETAAPKRTYDMMEGRVGRAISSASIEGLMGRAIP 1336
              : |:| |           | |     | :| ||             |||
Db       1554 LEQSEESGARNALVSEGDRGSPFQEEEGSALKR-------------SSAG---------- 1590

Qy       1337 PERHSPHHLKEQHHIRGSITQGIPRSYVEAQE 1368
                 :| || :   :: :  :|   |:   ::
Db       1591 ----APVHLGQGQFLKFTQREGDRESWSSGED 1618


RESULT 98
US-10-032-585-7665
; Sequence 7665, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
;  APPLICANT: Terry, Roemer D.
;  APPLICANT:  Bo, Jiang
;  APPLICANT:  Charles, Boone
;  APPLICANT:  Howard, Bussey
;  TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
;  FILE REFERENCE: 10182-005-999
;  CURRENT APPLICATION NUMBER: US/10/032,585
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 8000
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7665
;   LENGTH: 2093
;   TYPE: PRT
;   ORGANISM: Candida albicans
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (1501)..(1501)
;   OTHER INFORMATION: X=any amino acid
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (1519)..(1519)
;   OTHER INFORMATION: X=any amino acid
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (1526)..(1526)
;   OTHER INFORMATION: X=any amino acid
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (1576)..(1576)
;   OTHER INFORMATION: X=any amino acid
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (1604)..(1604)
;   OTHER INFORMATION: X=any amino acid
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (1608)..(1608)
;   OTHER INFORMATION: X=any amino acid
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (1610)..(1612)
;   OTHER INFORMATION: X=any amino acid
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (1615)..(1616)
;   OTHER INFORMATION: X=any amino acid
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (1621)..(1621)
;   OTHER INFORMATION: X=any amino acid
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (1624)..(1624)
;   OTHER INFORMATION: X=any amino acid
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (1647)..(1647)
;   OTHER INFORMATION: X=any amino acid
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (1668)..(1668)
;   OTHER INFORMATION: X=any amino acid
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (1691)..(1691)
;   OTHER INFORMATION: X=any amino acid
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (1693)..(1693)
;   OTHER INFORMATION: X=any amino acid
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (1695)..(1695)
;   OTHER INFORMATION: X=any amino acid
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (1712)..(1712)
;   OTHER INFORMATION: X=any amino acid
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (1729)..(1729)
;   OTHER INFORMATION: X=any amino acid
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (1785)..(1785)
;   OTHER INFORMATION: X=any amino acid
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (1787)..(1787)
;   OTHER INFORMATION: X=any amino acid
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (1790)..(1790)
;   OTHER INFORMATION: X=any amino acid
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (1841)..(1841)
;   OTHER INFORMATION: X=any amino acid
US-10-032-585-7665

  Query Match             2.7%;  Score 355;  DB 14;  Length 2093;
  Best Local Similarity   19.4%;  Pred. No. 9.1e-08;
  Matches  456;  Conservative  298;  Mismatches  963;  Indels  628;  Gaps  112;

Qy        491 GKSQQQQQQQQQ-----QQQQQQQQPMPRSSQE-EKDEKEKEKEAEKE-----EEKPEVE 539
              | |    ||         ||  | ||:  :::| |   |  : |  :|     |: |:
Db         35 GDDQSTSNQQDDFFSNIQQPDQPDQPIIETNEELETVHKSTDVELPQEIPLEQEQIPQQP 94

Qy        540 NDKEDLLKEKTDDTSGEDNDEKEAVASKGRKTANSQGRRKGRITRSMANEANSEEAITP- 598
               :|:::    ||  |  |:: :  ::  :   ||          :  | | |  : |
Db         95 QQQENVI----DDLFGTSNNDDDFFSNHNQDGLNS---------AEIETEENVEVPLQPD 141

Qy        599 -QQSAELASMELNESSRWTEE----EMETAKKGLLEHGRNWSAIARMVGSKTVSQCKNFY 653
               :|: :| | :   |    :|    | | :  :|::    :: |  : :::
Db        142 PEQNQDLESEKATISIDKKDEDPVLESLTTESDILDNDHESTSEAPGISTES-------- 193

Qy        654 FNYKKRQNLDEILQQHKLKMEKERNARRKKKKAPAAASEEAAFPPVVED-----EEMEAS 708
                    |: ::  :  | :  |    : :|  |:: :   |:  :     :| ||
Db        194 ----------EVFEEEAVVEEVQSTA----EDSPVDATQPSEPGPLSTEQINLVDEPEAG 239

Qy        709 GVSGNEEEMVEEAEALHASGNE------VPRG----ECSGP--------------ATVN 743
               :  | :   : :          ||       : |      | | |              |
Db        240 APNTEESQTESKLDQFFTESNEGDFFENIARDNVEQEVSNPSKTEADQKLEDLFAGADNE 299

Qy        744 NSSDTESIPSPHTEAAKDTGQNGPKPPAT-------LGADGPPPGPPTPPRRTSRAPIEP 796
              |  | : :  |    |:    |      ||       | |           ::  :  |
Db        300 NGEDLDFLQQPQHNVAE---VNSEAAVATESINLDLFGEDDTDFLDELNSKKQQQQEEEV 356

Qy        797 TPASEATGAPTPPPAPPSPSAPPPVVPKEEKEEETAAAPPVEEG-EEQKP---------- 845
              |:| :|  |     |  :      ::  ::  || :  | ||    | ||
Db        357 RPSSSSTTNPEDKLA--ALDLDDDLLLDDDFLEEGSVQPKVESSLPENKPVKHHHSYIPS 414

Qy        846 -PAAEELAVDTGKAEEPV-KSECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGSGRA 903
               |:| : :|      | |::  :  |:   |  ||      |    |   |    :  |
Db        415 GPSANQKSVYAPPQVSGVNKNDFIKNLEQSKKK-HDAYDFPDT----LLVNKSRPAARHA 469

Qy        904 TTAKSSGAPQDSDSSATCSADEVDEAEGGDKNRLLSPRPSLLTPTGDPRANASPQK---- 959
              ::   | |  | :|    :          |     |  |:  |||   : : :
Db        470 PSSNKYGPPSVSQTSIPSISSTSSLPPARDVKSSSSAPPAAAPPTGTSTSKVAKKSFFED 529

Qy        960 -PLD--LKQLKQRAAAIPPIQ-VTKVHEP---------PREDAAPTKPAPPAPPPPQ-NL 1005
               | |  : :  |:    |  |:: | ::| |           ::      : |||||| : |
Db        530 FPDDEEIPKPKRAGRAAAPVKAVNQIHSPQVLNQQVQQQQQQQQSQQNAPPAPPPSKVNK 589

Qy       1006 QPESDAPQQP--GSSPRGKSRSP-------------------APPADKEAF-------- 1035
              ||  : |  |  |:||  : ::|                 |||  :  ||
Db        590 QPPVN-PYMPKSGTSPILQQQTPAVHKSGIAAPNVNVPHANNYAPPLNSFAFPPRQPNQK 648

Qy       1036 ----AAEAQKLPGDPPCWTSGLPFPV-------------PPREVIKA---------SPHA 1069
                  : :  |  | ||  ||  |||                |  |            ||:
Db        649 TNDYSPQLQSSPAPPPPPTSLQPFPSIGQSQSAPVYHNRQPSNSINTNVNKVPSINSPYV 708
```

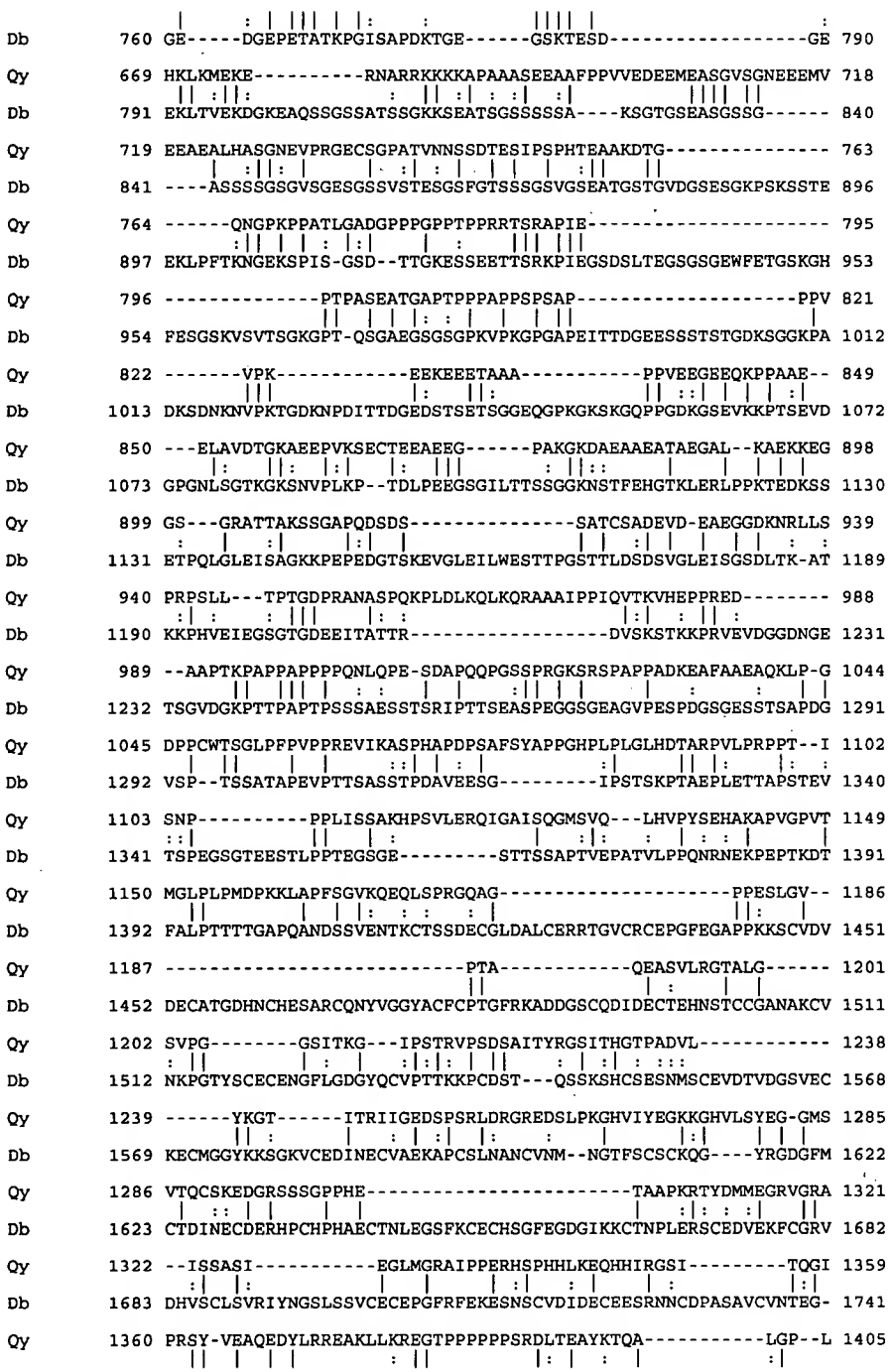
```
Qy       1070 PDPSAFSYAPPGHPLPLGLHDTARPVLPRPPTISNP--PPLISSAKHPSVLERQIGAISQ 1127
Db        709 --PNAGPYAPSTHQRS---HSRASSLVGGKGKEVNPYAPAL------ETVASGDINQIQQ 757

Qy       1128 GMSVQLHVPYSEHAKAPVGPVTMGLPLPMDPKKLAPFSGVKQEQLSPRGQAGPPESL--- 1184
Db        758 QQQQQQH----------LSPATAVLP----PARARGFSNAKSNIYKSAPKVVNPEQLVFR 803

Qy       1185 -----GVPTAQEASVL-------RGTALGSVPGGSITKGIPSTRVPSDSAITYRGSITHG 1232
Db        804 QYPIFSFSINSNCSVIIPNTGYNQGTSINIGPISNLLK-------DKDLYATFPGPLSKS 856

Qy       1233 TPADVLYKGTITRIIGEDSPSRLDRGREDSLPKGHVIYEGKKGHVLSYEGGMSVTQCSKE 1292
Db        857 KSK----KKDVEKWLESSSGYLRSLGKNDELVLNDILL-----HLIKHDGDVK----NQE 903

Qy       1293 DGRSSSGPPHETAAPKRTYDMME-GRVGRAISSASIEGLMGRAIPP----------ERHS 1341
Db        904 FLRSAC----SILNPNVNFDVNNLPTSGLAIGATNAHKLDNSGINTVFTLIQSGNLEKAL 959

Qy       1342 PHHLKEQHHIRGSITQGIPRSYVEAQEDYLRREAKLLKREGTPPPPPSRDLTEAYKTQA 1401
Db        960 EYCVSRGDWALALIISG-PENFNKIAAEYARNTFPFAKTNN--------------KVLH 1003

Qy       1402 LGPLKLKPAHEGLVATVKEAGRSIHEIPRE------------------ELRHTPELPLA 1442
Db       1004 IMPIILK-VLGGNVSSVIE---DLTAVPNEGEYANLHWREIISTIAICGSLKATEFLIEF 1059

Qy       1443 PRPLKEGSITQGTPLKYDTGASTTGSKKHDVRSLIGSPGRT----------------FP 1485
Db       1060 GKFLNQQGNFIGSEIAYIMAGLPFG---HQNFMMLASGHNSMYTEVYEYALNLHKAMHFP 1116

Qy       1486 PVHPLDVMADARALERACYEESLKSRP--GTASSSGGSIARGAPVIVPELGKPRQSPL-- 1541
Db       1117 HLLSLKL---KHALTLADYGLINESQRYIDHINSSIKTLGNKSPFVTPNLLHEFQNLIMR 1173

Qy       1542 -----TYEDHGAPFAGHLPRGSPVTMREPTPRLQE----GSLSSSKASQDRKLTSTPREI 1592
Db       1174 ITEVGSGDDQNNWFSGKISR---VNLDKIWGQIDKFIVGGDESKNGNNNDGNGT------ 1224

Qy       1593 AKSPHSTVPEHHPHPISPYEHLLRGVSGVDLYRSHIP-LAFDPTSIPRGIPLDAAAAYYL 1651
Db       1225 GNGSGSVFNKFSP-------SVSRNASSVNLHNYAQPSMIRQPSLLP----------YQ 1266

Qy       1652 PR-------------HLAPNPTYPHLYPPYLIRGYPDTAALEN---------RQTIINDY 1689
Db       1267 PQQQPQPQQQLLDQVHIERKPTTGFAPQPPPLVGHPSTTSVNKYSPSIKSSPRQAQQNKF 1326

Qy       1690 ----ITSQQMHHNTATAMAQRADMLRGLSPRESSLALNYAAGP-------RGIIDLSQVP 1738
Db       1327 EKYAPSNNSSHHN-------------LSLVEERSAVTSADGPEYPHHQHQQSINASAVP 1372

Qy       1739 HLPVLVPPTPGTPATAMDRLAYLPTAPQPFSSRHSSSPLSPGGPTHLTKPTTTSSSERER 1798
Db       1373 -----VPLPPPTPPVSMPQ--HVSRSPRSHQS-HQPPTLPPSHSHHVQQPS--------R 1416

Qy       1799 DRDRERDRDRREREKSI--LTSTTTVEHAPIWRPGTEQSSGSSGS-SGGGGGSSSRPASHS 1855
Db       1417 DRSPLATRIYPYSNSVGGQISTTSVGSIPSQIPLGRQTHGKQPSISSVISGDSI--AAVG 1474

Qy       1856 HAHQHSPISPRT-QDALQQRPSVLHNTGMKGIITAVEPSKPTVLRSTSTSSPVRPAATFP 1914
Db       1475 LGEQENVLPPSTGQTGKTATSEVNRNXEGYGFGGHYHHDQP----ETIXESPELRXLQQP 1530

Qy       1915 PATHCPLGGTLDGVYPTLMEPVLLPKEAPRVARPERPRADTGHAFLAKPPARSGLEPASS 1974
Db       1531 QSSEAEISKDISN--DVALDSAKIPEAS---QEPEEETDESGNVAAAPPPX-----PVA- 1579

Qy       1975 PSKGSEPRPLVPPVSGHATIARTPAKNLAPHHASPDPPAPPASASDPHREKTQSKPFSIQ 2034
Db       1580 --------PVAPP-------RKTKSSRSNPYAPSTDIGA-XSNAXSXXXQTXXGKP---- 1619

Qy       2035 ELELRSLGYHGSSYS-PEGV---EPVSPVSSPSLTHDKGLPKHLEELDKSHLEGELRPKQ 2090
```

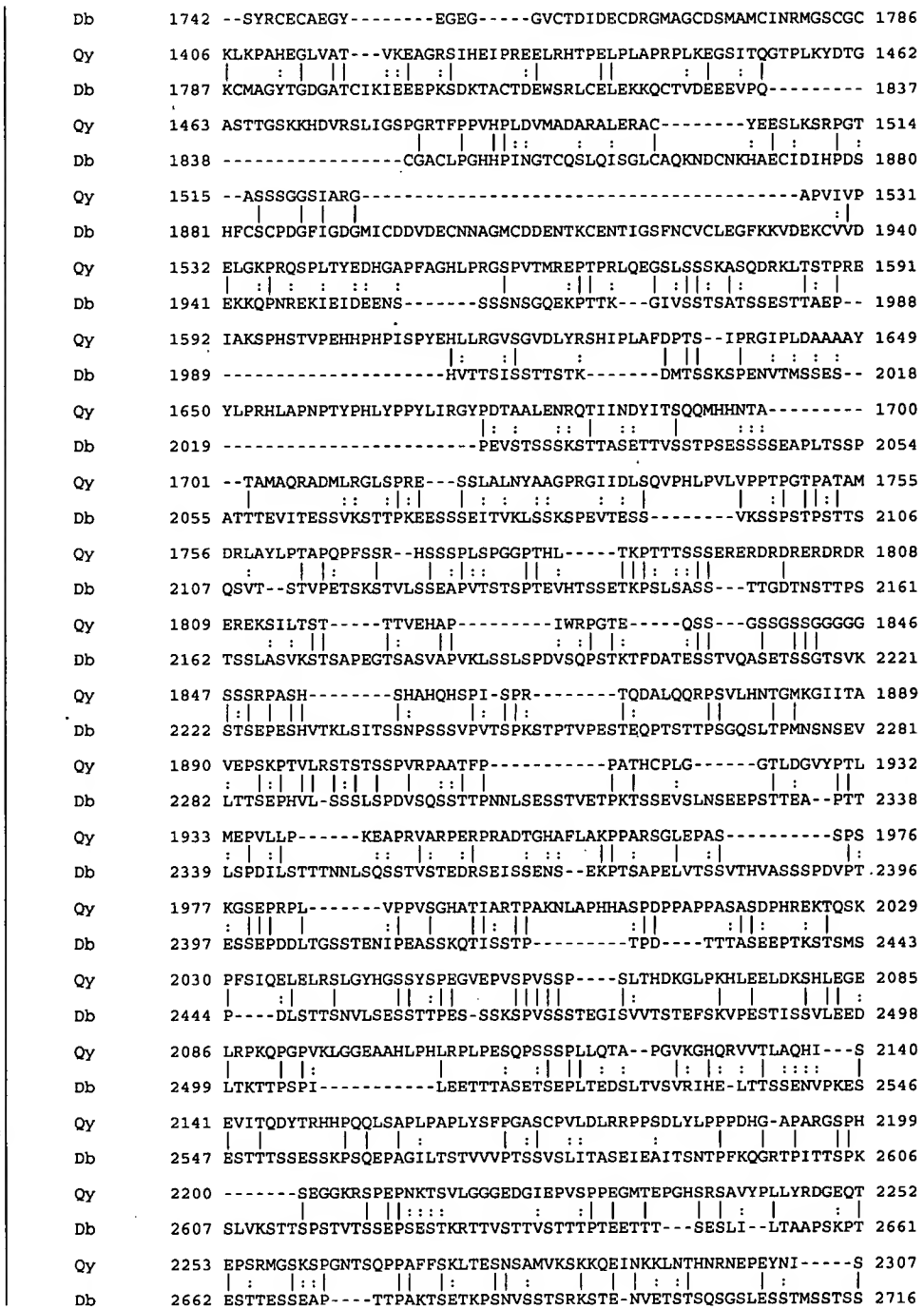
```
Db       1620 --SXRKXGSRTNRYGPPPGVGNKQPTIDVXPPSATNNTGNEDSISMFSYGXYQNESSPPL 1677

Qy       2091 PGPVKLGGEA-AHLPHLRPL------PESQPSSSPLLQTAPGVKGHQRVVTLAQHISEVI 2143
Db       1678 KQPSQFDQTAVASXPXPXPLQPQLAVPERVPTKNXANIDDSFDENSLAADTXTTYNNNMV 1737

Qy       2144 TQDYTRHHPQQLSAPL-PAPLYSFPGA--SCPVLDLRRPPSDLYLPPPDHGAPARGSPHS 2200
Db       1738 NKPYG-------SSPMGPTVATNGPGSVTSTPLI-LNQGSANMKL--------SNLSTIS 1781

Qy       2201 EGGKRSPEPNKTSVLGGGEDGIEPVSPPEGMTEP-----GHSR-----------SAVY-- 2242
Db       1782 VTGXGXGTXTGT---GGAFDGFPIPGSPDETTRPNSIFGGHTRGLFSSRLSESQSVLYQQ 1838

Qy       2243 ------------PLLYRDGEQTEPSRMGSKSPGNTSQPPAFFSKLTESNSAMVKSKKQEI 2290
Db       1839 YAXADDTVGDYIPIMEEDDEDDEDEQAKQQK-----------QKEKEAQEQELKRKQEQQ 1887

Qy       2291 NKKLNTHNRNEPEYNISQPGTEIFNMPAITGTGLMTYRSQAVQEHASTNMGLEAIIRKAL 2350
Db       1888 QQKAAAKNNN----NSGGGGGKFFSL--FGGGG-----NNKKQDN-------DAKVYKAH 1929

Qy       2351 MGK-----YDQ----W-------EES---------PPLSANAFNPL-------NASASLP 2378
Db       1930 LGQKNTFVYDEKLKRWIDKSKPLEEQLQANLPPPPPPSMKKKPNPVSTTTNSNNSSGPPP 1989

Qy       2379 AAMPITAADGRSDHTLTSPGGGGKAKVSGRPSSRKAKSPAPGLASGDRPPSVSSVHSEGD 2438
Db       1990 MVTPNPTPIPGGDLSTTATGG-------------ISTPIPASA----PPSRMS----KD 2027

Qy       2439 CNRRTPLTNRVWEDRPSSAGSTPFPYNPLIMRLQAGVMASPPPPGLPAGSGPLAGPHHAW 2498
Db       2028 STKPTPSIN-----RPKSSLATASGLDDLLSLTD----NNPPPNGRSRGVGGGATPGGSR 2078

Qy       2499 DEEPK 2503
Db       2079 RNKPK 2083
```

RESULT 99
US-10-369-493-5784
; Sequence 5784, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 5784
; LENGTH: 3507
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-5784

Query Match 2.7%; Score 354; DB 15; Length 3507;
Best Local Similarity 19.1%; Pred. No. 1.9e-07;
Matches 476; Conservative 288; Mismatches 879; Indels 848; Gaps 117;

```
Qy        550 TDDTSGEDNDEKEAVASKGRKTANSQGRRKGRITRSMANEANSEEAITPQQSAELASMEL 609
Db        715 TEEDEGEDEDLMEEGSGSWSTTINGTG-----ITGSPRSEGTIRVRIT----------TL 759

Qy        610 NESSRWTEEEMETA-KKGLLEHGRNWSAIARMVGSKTVSQCKNFYFNYKKRQNLDEILQQ 668
```





```
Qy       2308 QPGTEIFNMPAITGTGLMTYRSQAVQEHAST--------NMGLEAIIRKALMGKYDQWEE 2359
Db       2717 EPET---NAPAVTVSSEAS--STTLEENSSTSSPTSSEASVKLSSLFPESITSEAVTVSS 2771

Qy       2360 SPP----LSANAFNPLNASASLPAAMPITAADGRSDHTLTSPGGGGKAKVSGRPSSR--- 2412
Db       2772 RAPAEITMSSESHREISTVSSEPSEPEIPLSTTVSPNVVTA---------SSIPSEEPIL 2822

Qy       2413 ----KAKSPAPGLASGDRPPSVSSVHSEGDCNRRTPLTNRVWEDRPSSAGSTPF-PYNPL 2467
Db       2823 SSVTSSSTPRVRLITGTPDDLIVSVTVPSHGNRRQNIT----------ASSVPSNSTSPI 2872

Qy       2468 IMRLQAGVMASPPP----PGLPAGSGPLAGP 2494
Db       2873 ILPSESLTTPQPPPTTTTTAKPATTSGKRGP 2903
```

RESULT 100
US-10-058-124-18
; Sequence 18, Application US/10058124
; Publication No. US20030119058A1
; GENERAL INFORMATION:
; APPLICANT: Qvist, Per
; Bonde, Martin
; TITLE OF INVENTION: A Method for Assaying Collagen Fragments
; in Body Fluids, A Test Kit and Means for Carrying Out the
; Method and Use of the Method to Diagnose the Presence of
; Disorders Associated with the Metabolism of
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Darby & Darby PC
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/058,124
; FILING DATE: 29-Jan-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/570,573
; FILING DATE: 2002-MAY-12
; APPLICATION NUMBER: 08/187,319
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Gogoris, Adda C
; REGISTRATION NUMBER: 29,714
; REFERENCE/DOCKET NUMBER: 4305/08701
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-527-7700
; TELEFAX: 212-753-6237
; TELEX: 236687
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1341 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: COLLAGEN ALPHA 1 (I)
; SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-10-058-124-18

Query Match 2.7%; Score 351; DB 14; Length 1341;

Best Local Similarity 20.1%; Pred. No. 8.2e-08;
Matches 350; Conservative 106; Mismatches 573; Indels 714; Gaps 81;

```
Qy        722 EALHASGNEVPRGECSGPATVNNSSDTESIPSPHTEAAK-DTGQNGPKPPA-TLGADGPP 779
Db         77 ETKNCPGAEVPEGECC-PVCPDGSESPTDQETTGVEGPKGDTGPRGPRGPAGPPGRDGIP 135

Qy        780 -----PGPPTPPRRTSRAPIEPTPASEATGAPTPP----------PA 811
Db        136 GQPGLPGPPGPP----------GPPGPPGLGGNFAPQLYGYDEKSTGGISVPG 178

Qy        812 PPSPSAP-----PPVVPKEEKEEETAAAPPVEEGE--------EQKPPAAEELAVDTGKA 858
Db        179 PMGPSGPRGLPGPPGAPGPZGFZ----GPPGZPGZPGASGPMGPRGPPGPPGKBGBBGZA 234

Qy        859 EEPVKSECTEEAEEGPAKGKDAEAAEATAEGALKAEKKEGGSGRATTAKSSGAPQDSDSS 918
Db        235 GKPGR-----PGZRGPPGPZGARGLPGTA--GLPGMKGHRGFSGLBGAKGBAGPAG---- 283

Qy        919 ATCSADEVDEAEGGDKNRLLSPRPSLLTPTGDPRANASPQKPLDLKQLKQRAAAIPPIQV 978
Db        284 ----------PKGZPGS---------- 290

Qy        979 TKVHEPPREDAAPTKPAPPAPPPPQNLQPESDAPQQPGSSPRGKSRSPAPPADKEAFAAE 1038
Db        291 ------PGZBGAPGZMGPP----------GPKGNSGEPGAPGSK------ 318

Qy       1039 AQKLPGDPPCWTSGLPFPVPPREVIKASPHAPDPSAFSYAPPGHPLPLGLHDTARPVLPR 1098
Db        319 -----GD--TGAKGEPGPVGVQ-----GPPGPAGEEGKRGARGEPGPTG--------LPG 358

Qy       1099 PPTISNPPPLISSAKHPSVLERQIGAISQGMSVQLHVPYSEHAKAPVGPV-TMGLPLPMD 1157
Db        359 PPG----------ERG-GPGSRGF------PGADGVAGPKGPAGERGSPGPAG 394

Qy       1158 PKKLAPFSGVKQEQLSPRGQAGPPESLGVPTAQEASVLRGTALGSVPGGSITKGIPSTRV 1217
Db        395 PKG----------SP-GEAGRPGEAGLPGAKGLTGSPGS-----PGPDGKTGPPG--- 433

Qy       1218 PSDSAITYRGSITHGTPADVLYKGTITRIIGEDSPSRLDRGREDSL----PKGHVIYEGK 1273
Db        434 PAGQ---------DGRP----------GPPGPPGA-RGQAGVMGFPGPKGAAGEPGK 470

Qy       1274 KGHVLSYEGGM-----SVTQCSKEDGRSSSGPPHETAAPKRTYDMMEGRVGRAISSASIE 1328
Db        471 AG-----ERGVPGPPGAVGPAGKDGEAGAQGPP-GPAGPAGE----RGEQGPA-GSPGFQ 519

Qy       1329 GLMGRAIPP-ERHSPHHLKEQHHIRGSITQGIPRSYVEAQEDYLRREAKLLKREGTPPPP 1387
Db        520 GLPGPAGPPGEAGKPGE----------QGVPGDLGAPGPSGARGERGFPGERGVQGPP 567

Qy       1388 PPSRDLTEAYKTQALGPLKLK--PAHEGLVATVKEAGRSIHEIPREELRHTPELPLAPRP 1445
Db        568 GPA----------GPRGANGAPGNDGAKG---DAG----------APGAP-- 594

Qy       1446 LKEGSITQGTP----LKYDTGASTTGSKKHD-----VRSLIGSPGR------TFP--PVH 1488
Db        595 ---GS--QGAPGLQGMPGERGAAGLPGPKGDRGDAGPKGADGSPGKDGVRGLTGPIGPPG 649

Qy       1489 PLDVMADARALERACYEESLKSRPGTASSSGGSIARGAPVIVPELGKPRQSPLTYEDHGA 1548
Db        650 PAGAPGD----------KGESGPSGPAGPTGARGAPGDRGEPGPP--GPAGFA--GP 692

Qy       1549 PFAGHLPRGSPVTMREPTPRLQEGSLSSSKASQDRKLTSTPREIAKSPHSTVPEHHPHPI 1608
Db        693 PGAD----GQPGAKGEPGDAGAKGD----------AGPPGPAGPAGPPGPI 729

Qy       1609 ----SPYEHLLRGVSGVDLYRSHIPLAFDPTSIPRGIPLDAAAAYYLPRHLAPNPTYPHL 1664
Db        730 GNVGAPGAKGARGSAGP----------PGATGFPGAAGRVGP----PGPS-GNA 768

Qy       1665 YPPYLIRGYPDTAALENRQTIINDYITSQQMHHNTATAMAQRADML----------RGL 1713
```

```
Db        769 GPP----GPPGPAGKEGGKGPRGE-----------TGPAGRPGEVGPPGPPGPAGEKGS 812

Qy       1714 SPRESSLALNYAAGPRGIIDLSQVPHLPVLVPPT--PGTPATAMDRLAYLPTAPQPFSSR 1771

Db        813 PGADGPAGAPGTPGPQGIAGQRGVVGLPGQRGERGFPGLPGPSGE-----PGKQGPSGAS 867

Qy       1772 HSSSPLSPGGPTHLTKPTTTSSSERERDRDRERDRDREREKSILTSTTTVEHAPIWRPGT 1831

Db        868 GERGPPGPMGPPGLAGPPGESGRE----------------------------GAPGA 896

Qy       1832 EQSSGSSGSSGGGG-----GSSSRPASHSHAHQHSPISPRTQDALQQRPSVLHNTGMKGI 1886

Db        897 EGSPGRDGSPGAKGDRGETGPAGPPGAXGAXGAPGPVGP--------------AGKSGD 941

Qy       1887 ITAVEPSKPTVLRSTSTSSPVRPAATFPPATHCPLGGTLDGVYPTLMEPVLLPKEAPRVA 1946

Db        942 RGETGPAGP--------AGPVGPAGARGP----------------------------A 963

Qy       1947 RPERPRADTGH---------------AFLAKPPARSGLEPASSPSKGSEPRPLVPPVSGH 1991

Db        964 GPQGPRGDKGETGEQGDRGIKGHRGFSGLQGPPGPPGSPGEQGPSGASGP---------- 1013

Qy       1992 ATIARTPAKNLAPHHASPDPPAPPASASDPHREKTQSKPFSIQELELRSLGYHGSSYSPE 2051

Db       1014 -----------------AGPRGPPGSAGAPGKDGLNGLPGPI--------GPPGPRGRTG 1048

Qy       2052 GVEPVSPVSSPSLTHDKGLPKHLEELDKSHLEGELRPKQPGPVKLGGEAAHLPHLRPLPE 2111

Db       1049 DAGPVGPPGPPG----------------------PPGPPGPPSAGFDFSFLP------- 1078

Qy       2112 SQPSSSPLLQTAPGVKGHQ---------------------RVVTLAQHISEVITQDYTR 2149

Db       1079 -QP---------PQEKAHKGGRYYRADDANVVRDRDLEVDTTLKSLSQQIENIRSPEGXR 1128

Qy       2150 HHPQQLSAPLPAPLYSFPGASCPVLDLRRPPSDLYLPPPDHGAPARGSPHSEGGKRSPEP 2209

Db       1129 KNPAR--------------TC--RDLKMCHSDW----------------KSGEYWIDP 1154

Qy       2210 NKTSVLGGGEDGIEPVSPPEGMTEPGHSRSAVYPLLYRDGEQTEPSRMG-----SKSPGN 2264

Db       1155 NQ----GCNLDAIKVFCNME------TGETCVYP--------TQPSVAQKNWYISKNPKD 1196

Qy       2265 TSQPPAFFSKLTES----------------NSAMVKSKKQEINKKLNTHNRNEPEYNIS 2307

Db       1197 KDRHVWFGESMTDGFQFEYGGQGSDPADVAIQLTFLRLMSTEASQNITYHCKNSVAYMDQ 1256

Qy       2308 QPG 2310

Db       1257 QTG 1259
```

Search completed: April 15, 2004, 22:31:35
Job time : 531 secs